

In pursuit of systems

The study of functioning groups of molecules is an important frontier of biology at reductionist and holistic levels. Central to the long-term goals of scientific research, it brings its own challenges of infrastructure and evaluation.

What is the difference between a live cat and a dead one? One scientific answer is 'systems biology'. A dead cat is a collection of its component parts. A live cat is the emergent behaviour of the system incorporating those parts. There is certainly a vast distance to go before we can fully encompass such a system within scientific description. So how is systems biology already moving us towards the fullest possible description of a live cat?

By focusing on the behaviour of individual proteins and other biomolecules, much of what gives life its unique properties can be missed. To a systems biologist, the network of interactions formed by these components is more important than the molecules themselves. Properties such as robustness and evolvability, essential characteristics of life, then emerge from the topology of biological networks, independent of the constituents from which they are built.

Such a holistic view may sound dangerously soft-edged. Far from it. Systems biology couples the acquisition of comprehensive, high-definition data sets to the construction of quantitative models and computer simulations. Indeed, it is an explicit aim of both the Kyoto Encyclopedia of Genes and Genomes and the Alliance for Cellular Signaling to construct a fully functioning computer model of a cell.

The present state of experimental systems biology is both tantalizing and frustrating. To provide the level of detail required for us to know what is going on in a cell, microarray technologies will need to be faster and require smaller samples; ways to label and follow more biological molecules within a cell must be discovered; new spectroscopic tools to non-invasively measure multiple metabolite levels will need to be developed, and so on. *Nature* is committed to publishing studies that push back the technological frontier of what it is possible to know about important biological systems.

But technical wizardry and large data sets are only part of the systems-biology approach — a system is not fully understood until a quantitative model can be built. The role of modelling in biological research is controversial and can spark heated debates. What is clear,

though, is that the wealth of experimental data emerging from systems biology would be uninterpretable without detailed models against which they can be compared. Advances in modelling and simulation are thus no less important than data collection.

Every discipline generates community infrastructures, and systems biology is no exception. In the past five years, systems-biology institutes, departments and initiatives have been springing up across the globe. New journals have been launched, including The Institution of Electrical Engineers' *Systems Biology* and Nature Publishing Group's *Molecular Systems Biology*. The latter, an author-pays, online-only journal, is a joint venture with the European Molecular Biology Organization and went live last month.

The exchange of models between researchers is imperative, so a welcome development last month was the launch of BioModels (www.ebi.ac.uk/biomodels), a curated database for the deposition of biological models. BioModels has built on the success of Systems Biology Markup Language (SBML) in providing a format for the presentation of models, allowing them to be implemented on different software platforms. *Nature* journals and *Molecular Systems Biology* support submissions involving SBML.

It is hoped that BioModels will form the basis of a universally accepted repository that can do for systems biology what GenBank and the Protein Data Bank have done for genetics and structural biology. *Nature* applauds such efforts and will encourage authors of papers containing suitable models to contribute them to BioModels.

Systems biology presents an intellectual challenge to scientists and journal editors alike. Papers in this field document a highly multidisciplinary endeavour. Reviewers of such papers are very good at dissecting the aspects that fall within their sphere of expertise, but are less insightful beyond. So it falls to editors to weigh their frequently conflicting opinions in taking balanced and clear-sighted decisions. As a multidisciplinary journal, *Nature* welcomes the particular challenges that systems biology presents. ■

Fear and rambling at NASA

The US space agency needs an injection of rationality.

Only a confused space agency would consider shutting down the Voyager spacecraft as they approach the uncharted edge of the Solar System. Or cutting the basic research grants that provide the scientific basis for everything it does. Or cancelling satellites that make critical measurements of global climate change.

Last week a US National Academy of Sciences panel said that enough is enough, and called on NASA to reinstate some of its cancelled Earth-science projects (see page 9). NASA science chief Al Diaz denied that one of them, Glory, had even been cancelled. This highlights another disturbing trend: news of an impending cancellation is sent out or leaked, followed by a quick denial. Predictably, scientists reduced to chasing down rumours have turned fearful and angry.

NASA got into this mess through a variety of factors: an expensive initiative to send astronauts back to the Moon; uncertainty over how

much the space shuttle will cost to repair; an accounting system that figures in all the agency's overhead costs for the first time; moves by Congress to claim hundreds of millions of dollars from NASA's budget; and unprecedented congressional permission to move money between accounts. NASA managers who send out confused signals may not know themselves how much they can spend.

Into this muddle steps the new NASA administrator, Michael Griffin, who has already offered signs of hope. He, too, is impatient with the agency's accounting system and wants it fixed. He admits that not every project will survive NASA's change in direction, but at least wants decisions to be made rationally, meaning that the threatened Voyagers and the Hubble Space Telescope will get another hearing. The fiscal problems facing Griffin remain huge, but his willingness to address them honestly and directly are a good start. ■



Paper pushers

US government plans licence changes for foreigners
p4

Criminal errors

Survey of Iraq's weapons was thoroughly botched
p5



Wintry blast

Arctic ozone loss gives ecologists the shivers
p6

Knock on wood

Raised wildfire risk may follow US deforestation
p8

Pesticide results help China edge transgenic rice towards market

David Cyranoski

China is poised to commercialize genetically modified rice, possibly within a year. Data have just been published that advocate the rice's health benefits, but observers are concerned by the lack of long-term health and environmental studies, and by evidence that transgenic rice is already being sold illegally in the country.

China has four varieties of genetically modified rice in the final stages of field trials. Last week, Chinese and US researchers presented evidence that two of them — GM Xianyou 63 and GM II-Youming 86 — decrease the use of pesticides by 80% (J. Huang *et al.*, *Science* **308**, 688–690; 2005). The authors argue that this would dramatically benefit the health of farmers.

The authors' enthusiasm is in marked contrast to wary noises that have come from China in the past few years. No new transgenic food crop has been approved for commercial use there since 2000, and the government has taken every opportunity to emphasize the possible dangers.

This reflects China's 'no-risk' approach, claims a senior US-based plant scientist who advises the Chinese government on agricultural policy. The agriculture ministry has hesitated to introduce transgenic rice so as to avoid responsibility for any resulting health or environmental problems, he says, adding that positive results such as Huang's will be crucial in overcoming such fears.

But other observers have suggested that China has simply been taking advantage of biotechnology concerns in Europe and elsewhere. These enabled it to deny commercial approval to foreign companies such as Monsanto until its own domestic products could compete (see *Nature* **422**, 111–112; 2003).

The publication of the *Science* paper suggests that China believes it has now reached that stage, and is ready to be the first country to approve genetically modified rice. "Products from China's plant biotechnology industry could be an effective way to increase both

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Farmers in China may already be trading in illegal transgenic rice.

competitiveness internationally and rural incomes domestically," the authors write.

The results do sound impressive. GM Xianyou 63 and GM II-Youming 86 contain genes from the *Bacillus thuringiensis* bacterium and the cowpea, respectively, which make them resistant to two serious pests, stem borers and leaf rollers. Farmers in eight villages in the Hubei and Fujian provinces were given either conventional seeds or the transgenic varieties, all at the same price, and were then told to apply pesticides as needed.

There was a slight increase in productivity of about 6%. But it is the 80% drop in pesticide use that would bring both economic and health benefits, says the paper's lead author, Jikun Huang, an agricultural economist at the Chinese Academy of Sciences in Beijing.

Lesser poison

There are no studies showing the extent of pesticide-related illness in China, but according to Huang's unpublished estimates, around 50,000 farmers suffered pesticide poisoning each year in the 1990s, and around 1% of those died. Questionnaires submitted by the farmers in the latest survey showed that no households using the transgenic rice suffered symptoms of pesticide-related illness, whereas up to 8.3% of those using conventional rice did.

The drop in pesticide usage is not in itself a huge surprise, says Chris Leaver, a plant

biologist at the University of Oxford, UK, noting that similar advantages have been seen in the transgenic cotton that is already on sale. "It's nothing new except that it's in rice," he says.

But results such as Huang's could change perceptions about the costs and benefits of genetically modified food, says Robert Ziegler, who became director-general of the non-profit International Rice Research Institute in Los Baños, Philippines, earlier this year. "These hard data are consistent with the objectives that create these materials," he says. "It's not hype. It's real."

Selling genetically modified rice in China could lead to the acceptance of other crops, and other countries may follow suit, says Leaver: "This is just the tip of the iceberg." Huang confirms that scientists from Vietnam and Indonesia have contacted him about plans to introduce the pest-resistant rice. And progress is being made with Golden Rice, a transgenic strain that aims to combat malnutrition by producing extra pro-vitamin A (see J.A. Paine *et al.*, *Nature Biotechnol.* **23**, 482–487; 2005).

But it is not all good news. Some observers are concerned that the Chinese government may use Huang's results to justify approval of the transgenic strains, although few if any studies have been done on their long-term health or environmental effects. And they fear that lax regulation in China might cause problems. Transgenic rice already seems to be popping up illegally in markets in Hubei — GM Xianyou 63 is "probably" being sold there, admits Huang.

"Unregulated distribution would definitely happen within the country and across borders," says a Japanese plant biologist and advocate of genetic modification, who asked not to be named. He worries that unregulated distribution could lead to ecologically unbalanced agriculture. He argues that distribution of transgenic crops should come only after farmers have been educated about the technology and its possible risks.

M.S. YAMASHITA/CORBIS

Academics stress licence threat to US science

Geoff Brumfiel, Washington

Proposed changes to an obscure set of export rules could derail US research, say academic and industrial groups, who are now frantically trying to raise the alarm among scientists.

The modified rules would require academic researchers from countries including China and India to obtain a government licence before operating a wide range of lab equipment in the United States. In a 22 April letter to university department chairs, Judy Franz, executive officer of the American Physical Society, warned that the changes constitute a "potential threat to research". And this week, the National Academies are convening a special workshop to inform scientists of the proposed changes.

At issue is a set of rules governing the export of sensitive technologies. Known as the Export Administration Regulations, the rules are meant to limit the transfer of equipment that could advance the military might of 'countries of concern' — a list that includes China, India, Pakistan and Russia. The regulations also require researchers from these countries working with some items of equipment to obtain a licence from the US Department of Commerce.

Traditionally, universities have thought themselves exempt from the regulations. But a March 2004 report from the Department of Commerce's Office of Inspector General, an independent watchdog, argued that the regulations do apply to academic labs. The report also proposed expanding the criteria under which a licence would be required for using controlled equipment, and applying the rules

Dye lasers are one type of equipment that nationals from some countries may need a licence to use.

by country of birth rather than country of citizenship (see *Nature* **431**, 615; 2004).

The 45-page equipment list includes common lab apparatus such as lasers and sealed glove boxes for handling hazardous material. Getting a licence for each potential user would overwhelm lab supervisors, warns Dan Mote, president of the University of Maryland, who is scheduled to talk at a National Academies workshop. "This really is potentially devastating," he says. "It's quite conceivable that this would just bring work to a halt."

Industry is also concerned, according to Cynthia Johnson, director of government relations for Texas Instruments, a major US semiconductor manufacturer. Although industrial labs already have to comply with the

rules, the proposal to base the regulations on a researcher's country of birth rather than citizenship could alienate fresh talent, she says.

Department of Commerce officials stress that they are still far from making a final decision about how to modify the rules. "What we are doing is seeking input," says Peter Lichtenbaum, assistant secretary for export administration.

That is why it is important for researchers to weigh in with their objections, says Arthur Bienenstock, a physicist and dean of research and graduate policy at Stanford University in California. "What the Department of Commerce needs is an honest assessment of what it would mean if the inspector-general's rules were implemented," he says. The comment period closes on 27 May. ■

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Hubble rescue mission gears up despite shuttle setback

Tony Reichhardt, Washington

The space shuttle Discovery will have to wait at least another seven weeks before returning to duty, NASA decided last week. But preparations for an astronaut servicing mission to the Hubble Space Telescope will resume anyway — another sign that new NASA administrator Michael Griffin may reverse his predecessor's decision to abandon the telescope later in the decade.

NASA managers reluctantly revised Discovery's launch date from 22 May for several reasons, but primarily because of concerns about ice debris from the shuttle's giant fuel tank. The 2003 demise of Columbia, the last shuttle to fly, occurred after foam insulation falling from the tank shortly after launch gouged a hole in the shuttle's wing. NASA engineers consider foam the greatest danger, but a technical

review last week concluded that falling ice still poses a small risk. So heaters will probably be installed to prevent ice forming on the tank — a time-consuming job.

The next available launch window for the shuttle — determined by factors ranging from the International Space Station's orbit to lighting conditions for cameras to inspect the vehicle as it leaves the launch pad — opens on 13 July and lasts until 31 July.

Meanwhile, engineers at the agency's Goddard Space Flight Center in Maryland will resume planning for the shuttle servicing mission to Hubble that was suspended by former administrator Sean O'Keefe.

Griffin won't decide whether to approve the mission until after the shuttle's return, but said last week that he wanted to keep the option open. Based on new estimates of

when Hubble's batteries and gyroscopes will fail, project managers now think that the telescope can be serviced as late as 2010.

Griffin also said that, despite the delay in returning to flight, he considers the 2010 target for retiring the shuttle "very firm", and that NASA is looking for ways to complete the space station using other launch vehicles. This could mean that the schedule for conducting research on the station will slip.

For now, though, the orbiting laboratory will get what it needs from elsewhere: a Russian Progress vehicle will deliver supplies next month that the shuttle would have carried. And although the station's international partners are disappointed by the delay, said programme manager Bill Gerstenmaier, "they understand the decision". ■

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

An atmosphere of fear could have deterred former Iraqi weapons scientists from coming forward.

US bungled investigation into weapons research in Iraq

Jim Giles, London, and Geoff Brumfiel, Washington

Flawed interrogations of scientists in the aftermath of the Iraqi war hampered investigations of the country's weapons of mass destruction (WMD) programme and added to the risk that knowledge and technology would spread to neighbouring countries, US government advisers have concluded.

Iraqi researchers willing to cooperate with the United States were detained without cause while key WMD individuals were never targeted or even identified, according to new information released by the Iraq Survey Group (ISG). The ISG is a team of British, American and Australian inspectors charged with finding WMD after the fall of Saddam Hussein's regime. Coalition policies also created an atmosphere of fear that deterred potentially helpful researchers from talking and increased the likelihood that WMD scientists might defect to neighbouring states such as Iran or Syria, or find work with insurgent forces in Iraq, the authors said in a report released on 25 April.

"This report is being honest," says David Albright, a former United Nations weapons inspector and president of the Institute for Science and International Security in Washington DC. "The search was horribly mismanaged. The military treated scientists as common criminals."

Albright raised concerns soon after Baghdad fell to US forces, two years ago last month. He and other non-proliferation experts complained that Iraqi scientists were reluctant to cooperate for fear of being taken

prisoner and wanted protection from prosecution before they would talk (see *Nature* 423, 371; 2003). The ISG has now acknowledged these problems.

The public listing of the 55 most-wanted members of Hussein's regime is one tactic singled out for criticism. This left other Iraqis uncertain as to whether they were on a longer blacklist that included 300 names. As many Iraqis were at the time being detained for a wide range of reasons, "these factors combined to cause Iraq WMD participants at virtually all levels to attempt to remain undetected," the ISG says.

Those that were detained were often subject to bungled interrogation, the authors add. They found that army interrogators, who lacked technical training, glossed over key details when interviewing detainees and created inconsistent and sometimes error-filled intelligence briefs. WMD experts involved in the interviews usually only stayed in Baghdad for a month or two, often leaving just as they were gaining knowledge useful to the ISG investigation.

The ISG stresses that there is no evidence that defectors to neighbouring "hostile states" are actually working on WMD. But it adds that there are many reports of "Iraqis with general chemical or biological expertise helping insurgents to produce chemical and biological agents".

Around 100 of the 300 blacklisted individuals are still being detained. "As far as the WMD investigation is concerned, there is no further purpose in holding many of these individuals," the report concludes. ■

Rethink on review leaves researchers out in the cold

Geoff Brumfiel, Washington

US researchers could increasingly find that their grant applications are rejected without even being seen by funding-agency reviewers. The situation has arisen because the National Science Foundation (NSF) is asking universities to pre-screen some types of proposal in a bid to cut down its workload.

The NSF is responsible for funding the majority of US university-based, non-biomedical research. It has a sterling reputation among scientists for using peer review to determine who should receive grant money. But in the face of tight budget constraints, it is trying to shift some of that responsibility onto individual universities.

In the past few years, the NSF has placed limits on the number of applications that a single institution can submit. Those limits will now become increasingly common, according to Arden Bement, the agency's director. He says the measures are needed to control the number of proposals flooding in to his staff, and to boost the success rate of applications. He stresses that the new policy will affect only large facilities and collaborative grants. "This would not be for individual applications," he says.

But universities are starting to speak out about the proposals, warning that the changes are forcing them to become unwilling peer reviewers. Earlier this year, administrators at Princeton University, New Jersey, had to choose one of several proposals for a programme that funded international collaborations, according to Diane Jones, director of the university's office of government affairs. The proposals came from several disciplines and departments, making the choice far from straightforward. "Universities are not set up to do this kind of internal peer review," she says.

Fawwaz Ulaby, vice-president for research at the University of Michigan in Ann Arbor, adds that the rules particularly limit the opportunities for researchers at large universities. "They're doing this at the expense of fairness," he says.

But Bement says he has little choice in the face of a 2% budget cut this year. He adds that he still believes the agencies will receive the best proposals. "The universities will almost invariably put forward the ones of highest scientific merit," he says. ■

Free genome databases finally defeat Celera

Emma Marris, Washington

It's official: the genome wars are over. Celera Genomics, the company that raced against the public Human Genome Project to be the first to sequence our genetic blueprint, has announced that it will stop selling genomic information.

Celera was established in 1998 with the goal of commercializing access to genome-sequence information. Its bold plan to sequence the human genome in just a couple of years spurred on the public project, which made its data available free of charge. In June 2000, the two rivals declared the race a tie.

With so much genome data publicly available, the company realized as early as 2002 that the promised profits were not going to materialize. Its high-profile president Craig Venter, who had championed the sequencing effort, resigned to pursue other genomics interests.

Celera Genomics, based in Rockville, Maryland, will continue as a pharmaceutical company, working closely with its sibling Celera Diagnostics. Its drug-discovery efforts will build on some still-proprietary information, including a collection of single-nucleotide mutations from 39 humans and a chimpanzee. "It has become a different company in many ways," says Kathy Ordoñez, Celera's current president.

The last of the genomic information service, called the Celera Discovery System, will close by the end of June. All Celera's genomic data, including more recent mouse and rat sequences, will be made available in the public databases of the US National Center for Biotechnology Information in Bethesda, Maryland, in July.

The newly free human sequence will probably not be a great boon to researchers, as it is largely replicated by data from the Human Genome Project. But the rat and mouse sequences should be in more demand. "They will be quite useful because they are doing different strains," says Francis Collins, director of US National Human Genome Research Institute in Bethesda.

Collins, who led the Human Genome Project through the genome wars, calls the data release a "victory for the scientific community". Now he's speaking words of peace: "They have done a generous thing here and should be getting a lot of credit."

Arctic trends scrutinized as chilly winter destroys ozone

Quirin Schiermeier, Munich

The biggest ozone losses ever recorded over the Arctic have got scientists arguing over whether global warming is to blame. If there is a link, the spectre of an Arctic ozone hole looms.

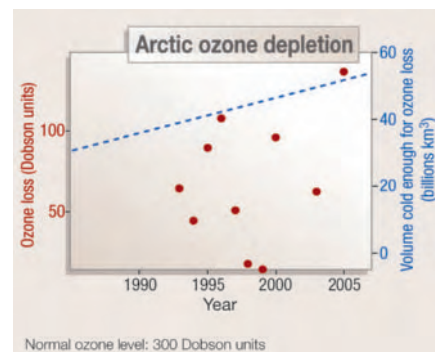
Researchers announced their preliminary results, from measurements taken between January and March, at last week's meeting of the European Geosciences Union in Vienna. Atmospheric scientist Markus Rex, of the Alfred Wegener Institute of Polar and Marine Research in Potsdam, Germany, and his colleagues reported that after this winter's unusually low Arctic temperatures, ozone losses are greater than any seen before (see graph).

The ozone layer shields the Earth from harmful ultraviolet radiation, but it is being eaten away by chlorofluorocarbons (CFCs). The existence of an ozone hole over the Antarctic was first reported exactly 20 years ago (Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* 315, 207–210; 1985); and production of CFCs was phased out by the Montreal Protocol just two years later. However, remnants of these long-lived chemicals are likely to linger in the atmosphere for at least another 50 years.

Ozone loss in the Arctic is less severe than in the colder Antarctic, and seems more dependent on temperature variation. Certain clouds in the stratosphere provide surfaces on which CFC decay products are converted into forms that destroy ozone — but the clouds only form at temperatures below -80°C .

The past Arctic winter was particularly cold and in some areas more than half of the ozone molecules were destroyed. By early spring, ozone-depleted air had drifted southwards through large parts of northern and central Europe.

Increasing levels of greenhouse gases are known to have a cooling effect on the stratosphere, because heat is locked near the sur-

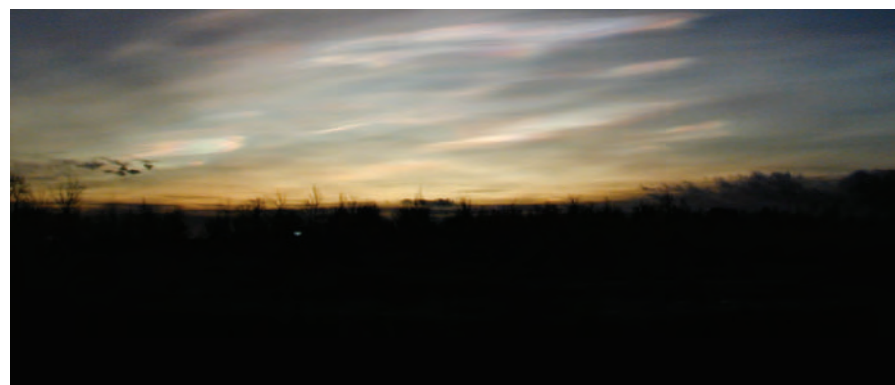


face. "Over the past 40 years, we have seen a fourfold increase in the area cold enough for polar stratospheric clouds," says Rex. He and his colleagues believe this year's losses are a sign that colder winters could lead to an Arctic ozone hole in the next two decades.

But other scientists are sceptical. "This has been a very interesting Arctic year, temperature-wise and ozone-wise," says Susan Solomon, an atmospheric chemist and co-chair of an Intergovernmental Panel on Climate Change working group. "But any attempt to link ozone loss to greenhouse gases is extremely speculative."

Solomon argues that the temperature drop this winter was far too large to be explained by greenhouse cooling, and that natural fluctuations in meteorological conditions probably far outweigh greenhouse effects. "You can't just look at radiation," she says, "the fact that the two preceding winters were very warm should give us pause."

Rex acknowledges that there are still many unknowns, although he says we should not have to wait too long for answers. His results are part of SCOUT-O3, a €15-million (\$20-million) collaboration involving over 100 scientists from 19 countries. The project aims at understanding the links between ozone, greenhouse gases and the climate system as a whole — it will run for four more years.



Long shadows: these polar clouds need low temperatures to form, but can wreak havoc on Arctic ozone.

SOURCE: M. REX, EUR. GEOSCI. UN. MEETING 2005

L. POOLE/NASA LANGLEY RESEARCH CENTER

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Fielding queries: studies of soybeans suggest plants will suffer from increased concentrations of near-surface ozone produced by global warming.

Hikes in surface ozone could suffocate crops

Jim Giles, London

If farmers don't adapt, changing environmental conditions could damage food production. That's the disturbing message from the first open-air simulations of Earth's climate in 2050, which suggest that negative effects of increasing surface ozone will outweigh any benefits triggered by warming.

Rising carbon-dioxide levels had previously led researchers to make rosy predictions about crop yields in coming decades — the gas promotes growth by increasing rates of photosynthesis, among other effects. But much of the work has been based on greenhouse trials and has ignored the impact of ozone concentrations near the ground. When researchers from the University of Illinois at Urbana-Champaign addressed this shortcoming in field trials of soya plants, they saw yields dive.

The result is a concern because the world's two largest soybean growers may face significant ozone increases, says crop scientist and study author Stephen Long. Urban air pollutants are expected to push up levels of near-surface ozone by at least 25% by the middle of this century, but rises in China — and in the US Midwest, where almost 300,000 square kilometres of soybeans are harvested annually — could be two to three times as great.

Ozone creates reactive molecules that destroy rubisco, an enzyme crucial for photosynthesis. It is also known to make leaves age faster. To investigate its effects, the Illinois team developed open-air experiments in which carbon dioxide and ozone were released from pipes that surrounded 16 plots of 200 square metres. By using wind sensors to control gas release, concentrations in the air over the plot were held close to the predicted 2050 levels over the three-year study.

The preliminary data, presented at a

Royal Society conference held over 26–27 April in London, suggests that yields will be cut by up to 10%. Previous forecasts of crop yields made by the Intergovernmental Panel on Climate Change (IPCC) should now be revised, warns Long. "No agrochemical company would use [greenhouse trials] to forecast how their chemical will affect crops, yet we use them to predict global food security," he complains.

The open-air studies revealed further threats that greenhouse work had not spotted,

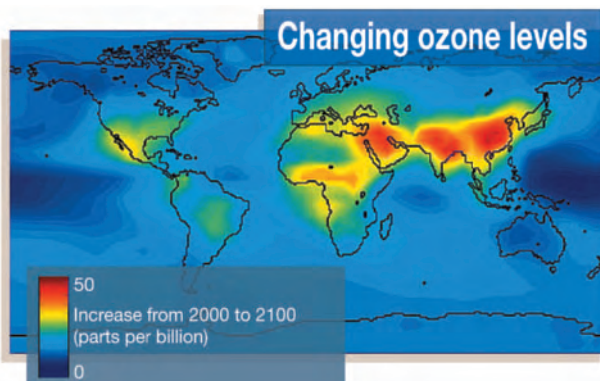
and Agricultural University in Taastrup, Denmark. He says that it is likely to be a particular problem in countries that are rapidly industrializing, such as India and China (see map).

At first, dropping yields would not be serious for developed regions such as Europe, where countries can import food and farmers are not pushed for the highest possible yields. But Porter points out that if other nations reduce production owing to ozone pollution, Europe may have to compensate.

Climate researchers are urgently working on ways to help farmers plan ahead. One approach is to combine climate simulations with software that models crop yields, to predict agricultural boom or bust.

Results from DEMETER, one of the most sophisticated models so far, are published in a series of papers in this month's *Tellus*. The EU-funded project forecasts growing-season weather using data from recent years. It runs multiple climate models and data sets at the same time — a technique known as ensemble modelling that has been widely adopted over the past few years as a way to cope with differing predictions from individual simulations. The outputs are then plugged into models of crop production. When tested for European wheat, DEMETER predictions averaged within 5.9% of the actual yield (Cantelaube, P. & Terres, J.-M. *Tellus A* 57, 476–487; 2005).

The approach is still in its infancy. But reliable crop forecasts could eventually benefit everyone from subsistence farmers to administrators. Farmers might, for example, choose not to plant a crop that is predicted to fail. And European Commission officials, who are showing an interest in the DEMETER model of wheat production, could adjust the payments they make to farmers to reflect which crops the continent is predicted to need in a given year.



including delayed maturation, which puts crops at risk of frost late in the season. And pests prospered. "The IPCC predicts that chewing insects will do worse in higher carbon dioxide," notes Long. "Unfortunately Japanese beetles like the environment. They live longer and produce more eggs."

The long view

All of the 22 soya varieties that Long tested showed a similar response to the conditions, suggesting that it might be hard to breed strains with better ozone tolerance. Within decades farmers may have to adapt by growing different crops, although without further studies it is unclear how other types of plant are likely to be affected.

"Ozone will be a threat," agrees John Porter, an ecologist at the Royal Veterinary

SOURCE: IPCC

Academy warning of bunker-buster risks leaves US unmoved

Washington A nuclear weapon designed to strike at targets deep underground would be likely to cause unacceptable civilian casualties, according to a study released by the US National Academy of Sciences.

Since 2002, the Bush administration has advocated the study and possible development of earth-penetrating nuclear weapons. Administration officials contend that such bombs would be the best way to strike at deeply buried targets — but other studies show that the bombs would spew out clouds of radioactive material that could harm civilians (see *Nature* **423**, 469; 2003).

The academy study shows that a nuclear weapon powerful enough to destroy a facility hundreds of metres below ground could kill up to a million people on the surface. It suggests that conventional weapons might be just as effective, without causing such devastating casualties.

Despite the new findings, administration officials are continuing to lobby for \$8.5 million for a study of the new weapon.

Wildfires add to growing problem of vanishing forest

Washington To cope with an increasing amount of land in need of reforestation, the US Forest Service should prioritize the issue and change the way it collects data. That's the conclusion of a report issued on 27 April by the Government Accountability Office.

Wildfires and other natural disturbances, such as insect infestations and diseases, have left 360,000 hectares in need of planting. If historically forested lands are not reforested, competing vegetation can take over, raising the risk of wildfires even further.

Forest Service officials acknowledge that changes need to be made, and say a new data tracking system is being developed. But they add that funding has not kept pace with the growing need for reforestation.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Burning issue: the loss of forested areas increases the risk of further wildfires.

G. BLEVINS/AP

NASA told to make space for Earth science

Washington Having watched with alarm as NASA axed or threatened to cut several key Earth-science missions as a result of budget shortfalls, a National Academy of Sciences panel has accused the agency of putting its decades-long Earth-science programme “at substantial risk”.

The committee, which is setting priorities for the next decade of Earth observations from space, recommended last week that two missions — Global Precipitation Measurement, and Atmospheric Soundings from Geostationary Orbit — be



reinstated immediately. It added that four more — Ocean Vector Winds, Landsat Data Continuity, Glory, and technology for a Wide-Swath Ocean Altimeter — be “urgently reconsidered”.

NASA's head of space science, Al Diaz, stated in a congressional hearing that the agency has “no intention of abandoning Earth science”, and said that not all the cancellations were final. For example, the agency intends to transfer some sensors to a new generation of weather satellites due to start flying in 2009, he explained.

Dash for cash raises nanotech patent fears

New York A “gold-rush mentality” towards patenting could inhibit the development of nanotechnology, warn US analysts.

The New York-based nanotech consultancy Lux Research analysed 1,084 nanotech patents, just under a third of those issued so far in the United States. The patents, which related to five different materials including carbon nanotubes and nanowires, often overlapped, leaving little space for new claims.

The report, issued late last month, adds that overly broad claims could be withdrawn by the courts, as has happened with numerous biotechnology patents.

But Stephen Maebius, a nanotech specialist in the Washington office of law firm Foley & Lardner, which co-authored the report, says that even though nanotech companies should be more careful about how they define claims, they should still aim to patent developments as soon as possible. “While some litigation is inevitable, those who have patents will have some leverage with which to avoid a self-destructive intellectual-property war,” he says.

Judges rule in favour of ‘saviour sibling’ babies

London Rejecting a challenge from an anti-abortion group, judges in the UK House of Lords ruled unanimously last week that the creation of babies through *in vitro* fertilization to help cure sick siblings is lawful. The decision marks the end of a lengthy battle in the courts about whether the nation's reproductive watchdog, the Human Fertilisation and Embryology Authority (HFEA), can allow couples to screen their embryos to find tissue matches for their seriously ill children.

Lawmakers backed the 2003 Court of

Appeal ruling that gave the HFEA the authority to grant Raj and Shahana Hashmi a licence to screen their embryos for this purpose. The Hashmis' son Zain suffers from the rare blood disorder thalassaemia major, and requires daily drugs and monthly blood transfusions. A transplant of stem cells from the umbilical cord of a ‘saviour sibling’ baby with the correct tissue match could cure his condition.

Advocates of reproductive rights have applauded the decision. “Families facing this critical situation will now have another treatment option,” says Laura Riley, director of the Progress Educational Trust, a UK charity that monitors issues relating to assisted reproduction and genetics.

Big bangs spark toad death explosion

Munich Speculation is running wild over harrowing scenes near a small pond in Hamburg, Germany. Three weeks ago passers-by first reported the sight of toads inflating like balloons and exploding. Since then, more than 1,000 toad corpses have been found, emptied of their innards, while scientists, the press and conspiracy theorists puzzle about the mass bursting.

Bacterial infections, environmental toxins, and a lethal fungus from a nearby race track have been blamed. But none of these has been confirmed by researchers at Hamburg's Institute for Hygiene and Environment, who are investigating cadavers and water samples. “We are facing a mystery”, says Janne Klöpfer, a spokesman for the institute.

Web focus: bird flu

A substantial collection of articles from *Nature*, news@nature.com and all relevant *Nature* journals has been created on our website.

♦ www.nature.com/nature/focus/avianflu

Meet the stripped down rat

Are wafers of silicon that support cells cultured from vital organs the future of drug testing and toxicology? Roxanne Khamsi talks to the pioneers creating model animals — and humans — on a chip.

For Michael Shuler, downsizing is not a dirty word. In the mid-1990s, his lab bench was cluttered with flasks connected by surgical tubing. Each flask contained cells from a different organ, suspended in nourishing fluids. Shuler's goal was to model how chemicals entering the body can be metabolized into more toxic forms.

Today, the apparatus has been shrunk to a tiny device that looks like something from the insides of your mobile phone. Shuler's lab is tidier, and his chips are more adept at simulating the way in which minute quantities of fluid carry chemicals between cells in our bodies. They may also point toxicologists towards a more humane future — reducing the number of animals required to establish the safety of a drug or other chemical. "We're not using this to replace animal studies," Shuler stresses. "We're trying to have more intelligent use of animals."

Shuler, a chemical engineer at Cornell University in Ithaca, New York, was converted to the chip-based approach in 1997, after Gregory Baxter arrived on the campus to take up a position developing nanofabrication technologies. "We got talking and saw that there were a lot of advantages of

Bare essentials: the 'rat on a chip' developed by Michael Shuler and Gregory Baxter.

adapting the system to fit on a silicon chip," Shuler recalls. Within a year, the pair had created a prototype 'animal on a chip' — a postage-stamp-sized silicon wafer that held cells from a rat's brain, heart and liver in different trenches linked by tiny fluid channels.

Channel hopping

Since then, Shuler has continued to refine his chips¹⁻⁴. The latest versions consist of two sheets of acrylic sandwiching a piece of silicon about 25 millimetres square containing fine custom-etched channels and troughs. The devices typically store around 100,000 liver cells, about 10,000 from another organ such as the lung, plus smaller numbers from tissues such as fat. "We can grow and keep cells on the chip

for weeks," says Shuler.

Liver cells are crucial because they are the body's workhorses for metabolizing drugs and other chemicals — in particular, they produce enzymes of the cytochrome P450 family, which oxidize a wide range of drugs. It is also important to mimic the way in which some chemicals are absorbed by the body's fat reserves — Shuler's team is experimenting with lipid gels⁵ in addition to fat cells.

Controlled by a micro-pump, pinkish fluid circulates from one chamber of cells to another. The nutrient liquid travels down channels about the size of a human hair, carrying drugs or other chemicals from one compartment to another. In these compartments, cells can cosy up to one another

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

"We're not using this to replace animal studies. We're trying to have more intelligent use of animals."

— Michael Shuler

much as they would in a living tissue, bathed in small quantities of fluid in which biologically active molecules can exert much greater effects than they would swilling around in a flask.

But Baxter and Shuler had to think carefully about the architecture of their chips, to simulate the ways in which cells are exposed to body fluids and dissolved chemicals in different organs. To mimic the rapid blood flow in the lung, for instance, the researchers initially made chambers containing multiple parallel channels. But this required more pressure, and disrupted the flow. So the researchers have since opted for larger chambers incorporating pillars on which lung cells can grow.

Proof-of-concept studies have highlighted the chips' value. Typical lab tests on single cell types, for example, struggle to reveal the lung toxicity of naphthalene, the main ingredient in mothballs. But Shuler's chips have no problem: the liver cells convert naphthalene into its toxic product naphthoquinone, a reactive compound which can then move to the 'lung' compartment where it damages the lung cells⁶.

On the move

In 2001, Baxter left Cornell to commercialize the technology at Hµrel in Beverly Hills, California, a company he co-founded. One of Hµrel's studies has used the technology to demonstrate the known action of tegafur, an anticancer drug that must be converted into an active form by enzymes in the liver. As expected, the drug had no effect on its target — colon cancer cells — unless it first ran through the liver-cell compartment to get switched on.

Shuler's group has now started using the chips to investigate multidrug-resistant cancer. Doctors have found that mixing medicines offers new hope for some patients, but the number of possible combinations is too great to be investigated in expensive clinical studies. With the chips, this is not a problem. "What we hope to do is screen treatments for disease that would not be screenable with animal or human studies," says Shuler.

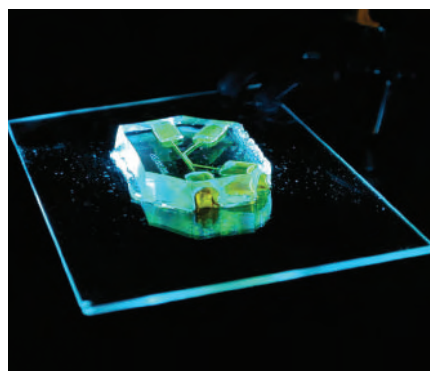
An extra bit of tweaking has even enabled Shuler's team to mimic the blood-brain barrier, which prevents many chemicals, including harmful ones, from entering the central nervous system. The endothelial cells that form this tight barrier quickly lose their special characteristics when cultured on their own. But by growing them alongside nerve-supporting cells called astrocytes on an ultrathin membrane, the researchers have managed to restore their protective function⁷.

The Cornell devices can run experiments for a couple of days, but problems then arise as bubbles of air start to form in the fluid. To try and get round this difficulty, Shuichi Takayama, a biomedical engineer at the

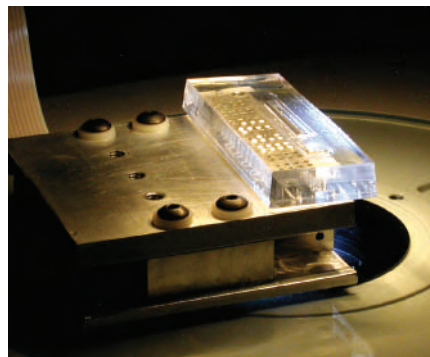
University of Michigan, Ann Arbor, is developing chips made from gas-permeable silicone gel. The trade-off is that the chips' contents slowly evaporate.

The pattern of flow from one compartment to another in Shuler's devices is predetermined for each application by the custom-etched pattern of channels, but in the Michigan version the fluid can be moved in different ways on the same chip. Takayama's group also has an unusual and novel approach to steering the fluids around.

The idea came from Wei Gu, an undergraduate student in Takayama's lab, who happened to read about a device called a



Shuichi Takayama's chips have been used to sort sperm (above). In the latest versions, fluid flow is controlled by an adapted Braille device (below).



refreshable Braille display while browsing in the university library. This hardware features a moving arrangement of pins that converts text from a computer screen into a form the blind can read by touch. Gu realized that he could apply the grid of vertically moving pins to open and close the fine channels on the pliable silicone chip. Following computer controls, the system pinches off flow much like a person stepping on a water hose — doing away with the need for more intricate pumps and valves⁸.

"Although many labs are working on microfluidic pumps, valves and even complete organ subsystems, devices that can be integrated onto a chip and be remotely controlled by a computer are rare," says Takayama. "Our goal is to not have to worry about microfluidics and worry about what to do with cells."

So far, Takayama's team has been cultur-

ing just one cell type on the chips, and has also been using them for tasks such as separating healthy from non-motile sperm. But like Shuler and Baxter, he aims to incorporate cells from different tissues and organs.

In addition to using samples from established cell lines for toxicological studies, Takayama and his colleagues plan to place undifferentiated stem cells on their chips. They then hope to mature these into nerve, bone and muscle cells, for example, by flowing the right signalling compounds through different paths, to study aspects of developmental biology.

Another dimension

Other researchers are trying to represent the three-dimensional structure of tissues more accurately, which can be important if cultured cells are to behave as they would in the body. Linda Griffith, a tissue engineer at the Massachusetts Institute of Technology, has focused her efforts on liver cells. At her lab in Cambridge, Massachusetts, researchers etch silicon to produce microscopic honeycomb-like scaffolds containing a multitude of channels and walls that cells can stick to⁹. These structures more accurately simulate the uptake of fluids in the liver, according to Griffith.

Hµrel might also experiment with similar chips in future. "It would be easy to modify the compartments to accommodate any three-dimensional tissue constructs," says Baxter. But for now, the company is concentrating on commercializing its current designs, the first of which could reach the market in about a year.

Either animal or human cells can be put on the chips. Shuler and Baxter began to incorporate the latter into some of their chips back in 2000. Toxicologists and drug developers typically work with rats because they need a living system that roughly matches human physiology. But in the end, notes Robert Freedman, chief executive of Hµrel, "a rat liver is not a human liver".

The new chips may not look much like rats or people. But if they can ensure that drugs reaching clinical trials have a better chance of being safe and effective than at present, they're likely to have a busy future. ■

Roxanne Khamsi is a London-based reporter for news@nature.com.

1. Sin, A., Baxter, G. T. & Shuler, M. L. *Proc. SPIE* **4560**, 98–101 (2003).
2. Park, T.-H. & Shuler, M. L. *Biotechnol. Prog.* **19**, 243–253 (2003).
3. Viravaidya, K. & Shuler, M. L. *Biotechnol. Prog.* **20**, 590–597 (2004).
4. Sin, A. *et al.* *Biotechnol. Prog.* **20**, 338–345 (2004).
5. Viravaidya, K., Jan, E. & Shuler, M. L. *Biotechnol. Bioeng.* **86**, 643–649 (2004).
6. Viravaidya, K., Sin, A. & Shuler, M. L. *Biotechnol. Prog.* **20**, 316–323 (2004).
7. Ma, S. H., Lepak, L. A., Hussain, R. J., Shain, W. & Shuler, M. L. *Lab Chip* **5**, 74–85 (2005).
8. Gu, W., Zhu, X., Futai, N., Cho, B. S. & Takayama, S. *Proc. Natl Acad. Sci. USA* **101**, 15861–15866 (2004).
9. Powers, M. J. *et al.* *Biotechnol. Bioeng.* **78**, 257–269 (2002).

Hotwire my heart

Growing numbers of people are being implanted with electronic devices that can automatically restart a failing heart. But have the risks and benefits been adequately assessed? Duncan Graham-Rowe investigates.

Cardiologist Michael Sweeney has vivid memories of the day he inadvertently set off an explosion inside a 79-year-old man's chest. It was 2001, and the patient had come into the Brigham and Women's Hospital in Boston for a routine check on his implantable cardioverter defibrillator, or ICD. Inserted under the pectoral muscles, an ICD can deliver a life-saving electric shock to a failing heart.

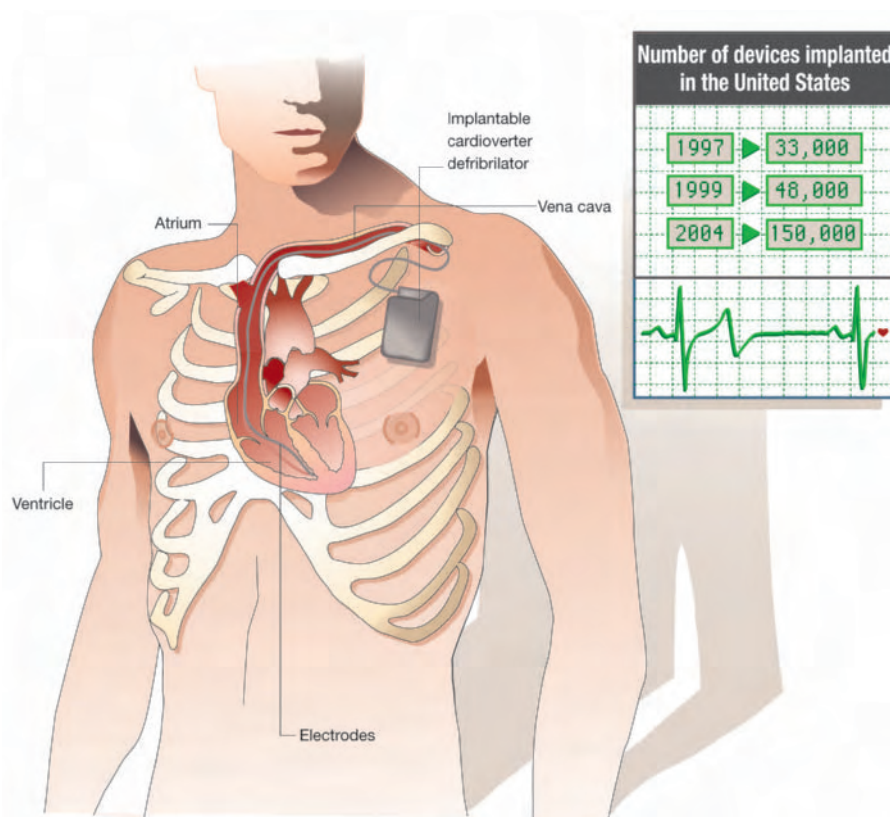
But when Sweeney attempted to set off a test shock from the device, there was a blue flash from the patient's chest and a loud popping sound. A short circuit in his ICD had ruptured its titanium casing¹. If the man hadn't been in hospital, says Sweeney, he almost certainly would have died.

To be fair, this same device, the man's fourth, had previously restarted his failing heart several times. But the incident illustrates that having one of these potentially life-saving implants is not risk-free. Although such catastrophic electrical failures are mercifully rare, the surgery required to install, maintain and replace ICDs exposes patients to the danger of infection. And with the devices becoming increasingly popular, experts are now beginning to wonder whether — for some of the patients now receiving them — the risks might outweigh the benefits.

"As defibrillators are being implanted in a larger number of people, the probability that each device will be required to save a person's life will go down," says Charles Swerdlow, a cardiologist at the Cedars-Sinai Medical Center in Los Angeles. Indeed, if your heart never actually fails, you will get all of the risks associated with having an ICD, and none of the benefits — apart, perhaps, from the peace of mind of knowing that the device is there, should you ever suffer a cardiac arrest.

Life saver

ICDs are extremely effective for people at high risk of heart failure. They apply shocks to the heart through conducting leads that are anchored within one of its chambers (see diagram, above right). They can either act as pacemakers, providing small pulses to maintain the heart's rhythm, or deliver much more powerful jolts. The latter can resynchronize the contractions of the heart's muscle if it has entered a chaotic or dangerously fast rhythm, or restart the heart if it stops beating.



Until recently, the only people who qualified for an ICD either suffered from potentially lethal abnormal heart rhythms, or had survived a previous cardiac arrest. But in recent years, official guidelines have extended their use to include many patients who are deemed likely to suffer heart failure. Over the past seven years, the number of people receiving ICDs in the United States has increased nearly fivefold.

Overall, this trend is thought to be saving lives. Indeed, in the largest ever ICD trial, published in January and involving more than 2,500 patients, the devices reduced mortality by 23% in people with 'moderate' heart failure, compared with those who were treated with drugs alone². On the basis of these results, the US Centers for Medicare and Medicaid Services expanded its criteria for treatment with an ICD. According to the leading ICD manufacturer, Medtronic of Minneapolis, 1.6 million Americans are now eligible for the devices. That's about five times the number of patients who have so far received one.

But some experts are sceptical of the need

to give implants to so many patients. Most of the trials of ICDs, they note, are sponsored by the manufacturers of the devices — which have a vested interest in seeing them used more widely. "I think inherently there is a conflict of interest," says Sweeney. "The companies are looking for a return on their investment."

Risk assessment

In particular, some cardiologists argue that more research should be done to distinguish between the patients at the highest risk of heart failure, and those who are unlikely ever to need a shock from an ICD. In the trial published earlier this year, for instance, only one in three of the devices delivered a shock over the five-year study². "Probably the large majority of those with ICDs will never have therapy," says Adrian Almquist, a cardiologist at the Minneapolis Heart Institute.

Some cardiologists argue that a series of simple, non-invasive tests could be used to identify the patients at highest risk of cardiac arrest. Certain features of an electro-

SOURCE: MEDTRONIC

cardiogram (ECG), for instance, may give a good indication of such risks.

In the trial published in January this year, too few patients had been subjected to these tests for definitive conclusions to be drawn². But in unpublished preliminary research, Morrison Hodges, also at the Minneapolis Heart Institute, has reanalysed ECG data from previous trials and found that they

could give an improved risk assessment for individual patients. Hodges now wants to conduct a multicentre trial in which thousands of ICD patients would be monitored over several years, in order to refine his predictive model. He approached Medtronic to see if it wanted to get involved, but the company wasn't interested. That's hardly surprising, says Hodges: "These tests are better at

inappropriately or to fail to be shocked when they need to be. The explosion that nearly killed Sweeney's 79-year-old patient was an extreme case, admittedly — the coiled lead of his device had worn through its insulation. But cardiologists fear that similar incidents may become more common in future. "The problem is that devices are getting smaller and smaller," says Almquist. This will make it harder to separate and adequately insulate the high-voltage circuits in them, he argues.

One worrying sign came in February, when Medtronic issued a warning that some of its Marquis model ICDs could develop a short circuit leading to the rapid discharge of the battery. The problem, a result of design changes made in miniaturization, would cause the device to emit an audible 'low battery' tone, or completely fail to work.

"We changed the manufacturing process and that problem does not occur any more," says Marshall Stanton, the company's vice-president and medical director, cardiac rhythm management. Only nine such failures have been reported, he adds, and the company's tests suggest that no more than 1.5% of 87,000 implanted Marquis devices are prone to the problem.

Hidden failure

But the true incidence of ICD failure remains unknown. A 2001 attempt to evaluate the problem concluded that mechanisms for reporting are inadequate³. In the United States and many other countries, there is no legal obligation for doctors to report a device failure and no central independent database to collate this information, says Almquist, who co-authored the study. John Camm, a cardiologist at St George's Hospital in London, suspects that only about one in ten ICD failures are reported.

"It is quite likely that there are dead patients whose devices failed," adds Sweeney. Under a different doctor, his 79-year-old patient could well have joined their number — Sweeney's clinic is unusual in that it runs checks in which patients are sedated and their devices are test-fired. "You have to turn over stones to find worms," he says.

Given the uncertainties, Sweeney and other experts argue that it's time to start turning over stones in a more systematic manner. If current trends continue, says Sweeney, we may get to the point where guidelines recommend ICDs for almost everyone with a possible risk of cardiac arrest. "And that is a very unsatisfactory position," he notes.

Duncan Graham-Rowe is a freelance journalist in Brighton, UK. He has a potentially lethal heart arrhythmia for which the only treatment is an ICD.

1. Sweeney, M. O. *J. Cardiovasc. Electrophysiol.* **12**, 1422–1424 (2001).
2. Bardy, G. H. *et al. N. Engl. J. Med.* **352**, 225–237 (2005).
3. Hauser, R. G. *et al. Pacing Clin. Electrophysiol.* **24**, 1046–1054 (2001).

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

An implantable defibrillator will restart a failing heart, but do so many patients need one?

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

identifying patients who don't need a defibrillator than those who do."

On cost grounds alone, there are good reasons to think carefully about who should be implanted with an ICD. Each device costs about US\$40,000, and every ICD patient requires lifelong care. The devices typically have to be replaced every five years or so, and they must regularly be monitored and serviced. Occasionally, parts such as the electrodes must be replaced. And each surgical intervention exposes the patient to the risk of infection. "This isn't like changing a part in your car," says Swerdlow.

And then there is the risk of electrical failures. These can cause patients to receive a shock



Non-invasive techniques could help decide who gets an ICD.

Guidelines reduce the risk of brain-scan shock

Responsibility for research is separate from a centre's duty of care to MRI volunteers.

Sir—Your News story “Brain-scan ethics come under the spotlight” (*Nature* 433, 185; 2005) raises an important issue, which the Wolfson Brain Imaging Centre (WBIC) has addressed over the past decade as part of a clinical research programme that includes normal volunteers.

This policy was designed to avoid the distressing scenario highlighted by a previous correspondent (“How volunteering for an MRI scan changed my life” *Nature* 434, 17; 2005) and to be consistent with UK guidelines for the safe operation of magnetic resonance imaging (MRI) scanners.

In brief, we separate the responsibility for the research itself, which properly lies with the investigator, from the duty of care owed by the WBIC. In addition, various forms of insurance are in place depending on the nature of the study, supplemented by a policy for non-negligent harm.

All volunteers are offered counselling, which includes discussion of what will happen should an abnormality be detected on an MRI study. Our information/consent

forms include the following paragraph:

“There is a chance of less than 1:100 that your MR scan may show a significant abnormality of which you are unaware. In such circumstances ... you will be referred to the appropriate specialist in consultation with your general practitioner, if that is what you would like. Such early detection has the benefit of starting treatment early but, in a small number of cases, may have implications for future employment and insurance.”

A record is kept of MR exposure for each individual. All structural MRI studies are reviewed by a consultant neuroradiologist and a confidential report is generated, which is not included within the normal hospital information system.

Any significant abnormality is then discussed with the clinical director of the WBIC, who is a consultant neurosurgeon in active clinical practice.

If the abnormality may affect the validity of the research, the principal investigator is informed that it is no

longer appropriate to include that individual in their study. The researcher is not told why and plays no further part in the process.

The volunteer is informed that there may be an abnormality and a full clinical MRI investigation is arranged if necessary. Volunteers are reassured that no communication will be made with their family doctor unless they so wish. This is to avoid details entering the medical records that might be used at some future date by life insurers and so on.

Finally, care has to be taken to ascertain that subjects are genuinely healthy volunteers. A very small number of volunteers are serial attenders at imaging centres worldwide. These people appear to use the system as a way of checking on their pathology, which they do not always declare at the time of screening.

John D. Pickard, Jonathan H. Gillard

Wolfson Brain Imaging Centre,
University of Cambridge, Box 65,
Addenbrooke's Hospital,
Cambridge CB2 2QQ, UK

DNA barcoding a useful tool for taxonomists

Sir—The Consortium for the Barcode of Life (CBOL; see www.barcoding.si.edu) is an international initiative of natural history museums, herbaria, other biodiversity research organizations, governmental organizations and private companies which wish to promote the development and use of DNA barcoding.

CBOL is in complete agreement with the major point raised by M. C. Ebach and C. Holdrege in Correspondence, that “DNA barcoding is no substitute for taxonomy” (*Nature* 434, 697; 2005).

CBOL views barcoding as a useful tool for taxonomists and a cost-effective system with which non-specialists, such as border inspectors, can assign unidentified specimens to known species. In both cases, CBOL views barcoding as part of taxonomy and rejects the idea that DNA taxonomy will replace the practice of taxonomy based on diverse character sets.

Taxonomists have begun using DNA barcodes in three ways. First, barcoding can be used as a ‘triage’ tool for sorting new collections into units based on barcode sequences, of which some will belong to known species and others will be new to science. In CBOL's view, only expert taxonomists can resolve the relationship

between new barcode-based clusters and species.

Second, DNA barcodes can also help assign specimens to known species in those cases where morphologic features are missing (in the case of immature, partial or damaged specimens) or misleading (as in sexually dimorphic species). Third, barcodes can also be used as a supplement to other taxonomic datasets in the process of delimiting species boundaries.

Ebach and Holdrege are correct in stating “DNA barcoding generates information, not knowledge”. CBOL believes that this information can make systematists and the consumers of taxonomic information more knowledgeable. Therein lies its potential value.

David E. Schindel, Scott E. Miller

Consortium for the Barcode of Life, National
Museum of Natural History, Smithsonian
Institution, PO Box 37012, MRC-105,
Washington, DC 20013-7012, USA

Debate to be had on best practice in mentoring

Sir—I much appreciated your Editorial “Joys of (top-notch) supervision” (*Nature* 434, 421; 2005). Having supervised 35 successful PhD students—and resisting

the temptation to write a book on the subject—I believe that there is a debate to be had on good practice in student project mentoring and the problems of identifying and avoiding poor supervision.

For example, universities have procedures for recognizing academic staff who are qualified to supervise research students, but, in my experience, they lack adequate processes for dealing with poor supervisors.

To focus on one particular point, that of co-authorship of publications with the supervisor: the supervisor should indeed have contributed to the scholarship substantially beyond the baseline of normal supervisory support.

However, practice varies between disciplines, and in certain areas of the physical and biomedical sciences it is the norm for the supervisor's name to appear on any publication arising from a student's research.

Alan Lord

University College London,
Gower Street,
London WC1E 6BT, UK

correspondence

Contributions to Correspondence may be submitted to corres@nature.com. They should be no more than 500 words in length, and ideally shorter. Published contributions are edited.

Gödel's universe

The legacy of one of the greatest thinkers of the twentieth century.

Incompleteness: The Proof and Paradox of Kurt Gödel

by Rebecca Goldstein

W. W. Norton: 2005. 288 pp. \$22.95.

To be published in the UK by Weidenfeld & Nicolson.

A World Without Time: The Forgotten Legacy of Gödel and Einstein

by Pallo Yourgrau

Basic Books/Allen Lane: 2005. 224 pp. \$24/£20

Martin Davis

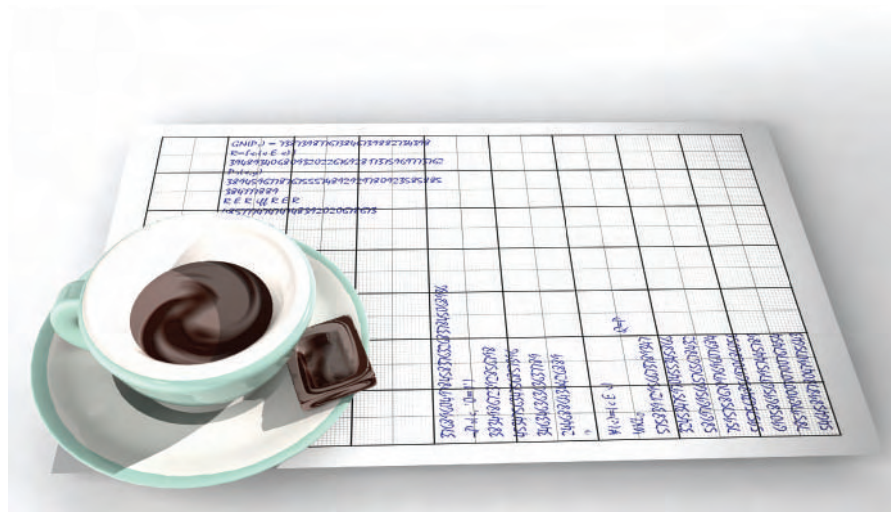
When the US magazine *Time* announced its selection of the “twenty greatest thinkers and scientists of the twentieth century”, readers would hardly have been surprised to find that Albert Einstein was included. But how many readers, seeing the name Kurt Gödel on the list, would have had any idea who he was, or what he had done to deserve this accolade? Both these well-written books tell the story of the life of this strange and tormented man, and explain some of his accomplishments.

Born in 1906 to a German-speaking family in Brno (which is today in the Czech Republic), Gödel was educated at university in Vienna, at first studying physics but soon finding that mathematics was his true métier. He was particularly attracted by the rigour of mathematical methods and the certainty of mathematical truth, so the controversies over the validity of these methods that raged during the 1920s fascinated him.

A great battle for the soul of mathematics was being fought between David Hilbert, arguably the greatest mathematician of his time, and the dissident upstart L. E. J. Brouwer, who vehemently denied the validity of some of the most cherished methods in the mathematician's armoury. Hilbert fumed that Brouwer and his allies “seek to provide a foundation for mathematics by pitching overboard whatever discomferts them”.

Hilbert's radical remedy was to create an entirely new branch of mathematics, dubbed metamathematics, which would apply mathematical methods to mathematics itself. In order to achieve this, it was first necessary to present mathematics as a ‘formula game’ in which the propositions of mathematics were represented as mere assemblages of symbols, and the methods of inference that led from axioms to theorems were presented as transparent rules for manipulating symbols.

Previous work by Gottlob Frege and by the mighty team of Alfred North Whitehead and Bertrand Russell had shown how to develop all of mathematics in this manner.



Hilbert and his co-workers were trying to prove that no contradiction would ever arise in the formula game, restricting themselves to methods even more severe than those countenanced by Brouwer.

The problem that Gödel chose for his doctoral dissertation was posed in a little textbook on formal logic published in 1928 by Hilbert and one of his students: to prove that the methods of inference used in the formula game encompass all possible logical reasoning.

After a quick success, Gödel chose for his *Habilitation* (the second doctorate required in the German-speaking world for an academic career) a central problem of Hilbert's metamathematics: the consistency of the arithmetic of real numbers. Hilbert's followers were already making good progress, or so it seemed, on the easier problem of proving the consistency of the arithmetic of whole numbers. So Gödel proposed to prove the consistency of the arithmetic of real numbers relative to that of whole numbers; that is, to show that any inconsistency with real numbers would in turn produce an inconsistency with whole numbers.

It was while Gödel was working on this problem that he came to an amazing conclusion that transformed him from a supporter of Hilbert's metamathematics into its nemesis. He saw that no version of the formula game could encompass all truths, even about whole numbers. In the published version of his discovery, Gödel introduced a code by which the expressions of the formula game were each represented by a single whole number. Thus, certain statements about numbers could be seen by someone privy to the code as making assertions about the formula game as well. Gödel showed how to construct a statement P that, in this manner,

asserted that a certain other statement, Q, could not be proved using the rules. Then he showed how to make $Q = P$, so that the statement P actually asserted that P itself is unprovable. From this it follows that P must be true and therefore, because of what it asserts, unprovable.

This necessary incompleteness of the formula game was the first blow to Hilbert's work. A few weeks later, Gödel delivered the knockout punch by proving that Hilbert's cherished goal of proving the consistency of the formula game through restricted methods was doomed to failure.

Rebecca Goldstein is a distinguished author of several novels that reflect her background as a philosopher. She writes with a light touch that readers are sure to enjoy, although they should be warned that she is not always entirely accurate. Assuming that the particular code that Gödel used for his proof had special properties needed to make it work, she argues that “superhuman efforts” were needed to settle on that code. Actually, any other simple code would work as well. She properly emphasizes Gödel's platonist position, according to which numbers and sets have an objective existence, but she doesn't see that he only adopted this view years after his sensational proof of incompleteness, so she misses the important fact that Gödel set out not to bury Hilbert's programme but to contribute to it. Finally, she quotes Hilbert as saying: “Mathematics is a game played according to certain simple rules with meaningless marks on paper.” It seems doubtful that Hilbert ever said this. For him, the formula game was not the be-all and end-all of mathematics, but only a means to an end, securing mathematics from contradiction.

Palle Yourgrau is a philosopher who has

written extensively about Gödel's contribution to Einstein's theory of relativity. Yourgrau and Goldstein both write about the friendship that developed between these two very different scholars. Gödel was a shy, reclusive man with a tendency to paranoia, whereas Einstein was worldly and outgoing. Nevertheless, the two were seen each day walking together to and from their offices at the Institute for Advanced Study in Princeton, obviously enjoying one another's company.

The equations of general relativity, having superseded Newton's account of gravitation, provide one of the great successes of

twentieth-century physics. In 1949, Gödel discovered unexpected solutions to these equations corresponding to universes in which no universal temporal ordering is possible. A hypothetical inhabitant of such a universe could, in principle, travel to his own past. Yourgrau argues that because time fails to exist in these Gödel universes, and because the differences between such universes and our own are accidental, time can't exist in our world either. I doubt that many readers would be convinced by this argument.

When his wife was hospitalized, Gödel literally starved himself to death, unwilling to eat anything not prepared by her. Referring

to his sad end, Goldstein makes the untoward suggestion that he might have imagined that he was living in an actual Gödel universe in which he could look forward to an eternal recurrence, reliving his life over and over again. A staunch believer in an afterlife, Gödel would hardly have sought such a fate. In any case, weird as the Gödel universes are, such eternal recurrence is not one of their properties. ■

Martin Davis is professor emeritus, New York University, and is a visiting scholar at the University of California, Berkeley. He can be contacted at 3360 Dwight Way, Berkeley, California 94704-2523, USA.

The dark side of astronomy

Empire of the Stars: Friendship, Obsession and Betrayal in the Quest for Black Holes

by Arthur I. Miller

Little, Brown/Houghton Mifflin: 2005.

416 pp. £16.99/\$26

Thanu Padmanabhan

Science, unfortunately, is done by humans. This makes scientific debates subject to the usual follies of human interaction, prejudices and lack of objectivity. One classic example in the modern history of astrophysics is the dispute between Arthur Eddington and Subrahmanyan Chandrasekhar (also known as Chandra) regarding the fate of stars with masses above a critical value. In 1930, at the age of just 19, Chandra made the pioneering discovery that such objects should continue to collapse to a singularity, a mathematical point of infinite density. Eddington, the leading authority in astrophysics at the time, ridiculed this conclusion. And thereby hangs a tale.

The fate of such massive objects is an extremely important question in astrophysics. Chandra was predicting the existence of black holes, although this idea was years away from reaching maturity. He reached his viewpoint by combining the laws of special relativity with those of quantum theory as applied to particles such as electrons when they are in a state in which they are said to be relativistically degenerate. Chandra's calculations, related to the pressure of the relativistically degenerate gas, led him to the conclusion that gravity will inexorably crush a sufficiently massive body to a point.

If Eddington had merely expressed discomfort at matter collapsing to a point-like singularity of infinite density, that would have been fair enough. Even today we do not understand what actually happens to matter that falls into a black hole, and those

of us who have thought about it are uncomfortable with the notion of a singularity. But Eddington questioned Chandra's logic and his calculation, and raised irrelevant objections. Later work completely vindicated Chandra, but the early humiliation left a deep scar. We can only speculate about what would have happened if Eddington had taken the scientifically correct stand and helped to develop Chandra's ideas further.

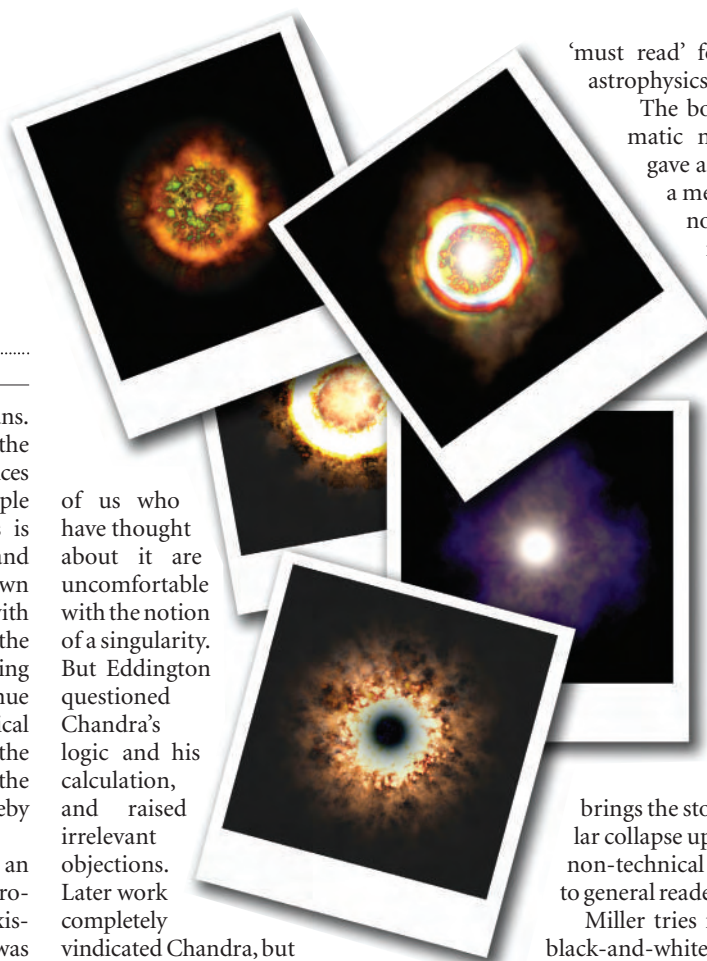
Historian Arthur Miller tells this story in a lively and exciting narrative in *Empire of the Stars*. The book is so beautifully written that I read it in one sitting. The scientific issues, as well as the personalities of everyone involved in the debate, have been described in picturesque detail without either mincing words or making unsubstantiated claims. Wherever necessary, Miller is careful to give references to the sources from which he has gathered the information, making the work fairly authoritative. In short, this book is a

'must read' for anyone interested in astrophysics or the history of science.

The book begins with the dramatic moment when Chandra gave a seminar on his results at a meeting of the Royal Astronomical Society in London in January 1935, and describes how Eddington made mincemeat of it. It then goes backwards in time, caricaturing both of the dramatis personae, Chandra and Eddington. Having put the pieces together, Miller goes on to describe how this particular episode left its mark on Chandra. As well as the story of Chandra's encounter with Eddington, Miller tells of the progress made in this subject over the years, and an appendix

brings the story of supernova and stellar collapse up-to-date. The language is non-technical and the book is accessible to general readers.

Miller tries not to present people in black-and-white terms. He succeeds to a remarkable extent, and much of the discussion is appropriately candid. For example, he doesn't shy away from mentioning the racial prejudices that existed in Chandra's time. And he has no hesitation in referring to Eddington's visits to Chandra in the weeks before the fateful seminar — to check on what he was doing — as "sheer mean-spirited duplicity". On the other hand, Miller explicitly points out Chandra's own prejudices and underhand dealings. In a letter to Léon Rosenfeld, Chandra explains how he added to one of his papers a reference to James Jeans, Eddington's arch-rival and the likeliest reviewer of the paper, in order to get the paper published. Miller quotes from Chandra's letter: "It is really sickening —



these underhand methods, but what can one do?" Similarly, Miller points out that Chandra continued to complain bitterly that leading physicists never supported him against Eddington, despite the fact that Paul Dirac, Rudolf Peierls and Maurice Price wrote an important paper that backed him up.

On the sidewalks of the narration, one also picks up some interesting snippets, such as how the Nobel laureate Raman declared that there will be no astrophysicists within miles of Bangalore; or how Chandra decided to spend some time in America because of the "underhanded dealings going on in Indian scientific circles"; and why Robert Oppenheimer had a successful collaboration with Richard Tolman.

In summary, this is an entertaining and illuminating book about a key issue in contemporary astrophysics. The author is to be congratulated on producing an authoritative description in a manner that is so thoroughly enthralling.

Thanu Padmanabhan is at Pune University, Ganeshkhind, Pune 411 007, India.

Destroying the zombic hunch

Sweet Dreams: Philosophical Obstacles to a Science of Consciousness

by Daniel C. Dennett

MIT Press: 2005. 216 pp, \$28, £18.95

Susan Blackmore

Has the devil lost his horns? Dan Dennett has been demonized because of his tough materialist stance on consciousness and his claim that we are conscious machines with nobody at home inside. But now he updates his theory and, I think, presents a softer version of it.

Sweet Dreams is a collection of essays and lectures written between 1999 and 2005 in which Dennett tries to freeze time and present a 'best' version of his evolving ideas. There are problems with this, especially as some chunks are somewhat disconcertingly repeated from one chapter to the next, but the overall picture gives a good idea of where Dennett's thinking has been going.

He describes his task as being to explain away qualia-based intuitive objections to materialism, and this he does using some of his favourite examples. One is the ever-popular philosopher's zombie, an imaginary creature who looks, acts and speaks like a normal person but has no subjective

experience or qualia. It's easy to imagine a zombie — or at least to think you are imagining one — says Dennett, and he calls this "falling for the zombic hunch", which traps people into believing that consciousness is separate from brain function. Dennett has tried to murder the zombie before, explaining how people fail to follow the rules when they think they are imagining one, but now with the concept of the zombic hunch he explores the damage done by this false intuition. Don't worry, he says, if you are patient and open minded it will pass, or mutate into a less virulent form. Just as we still feel as though the Earth stands still, in the future people may still feel the zombic hunch, but they won't believe it. They will know that mechanistic theories of consciousness do the whole job, so we don't need the concept of qualia.

Demystifying consciousness is Dennett's forte, and is probably the main reason for his status as the devil — after all, people like mysteries. Some even think that if consciousness is explained they will be diminished as people, being turned into mere 'things'. They don't like the idea that there is no one at home inside their head, no audience

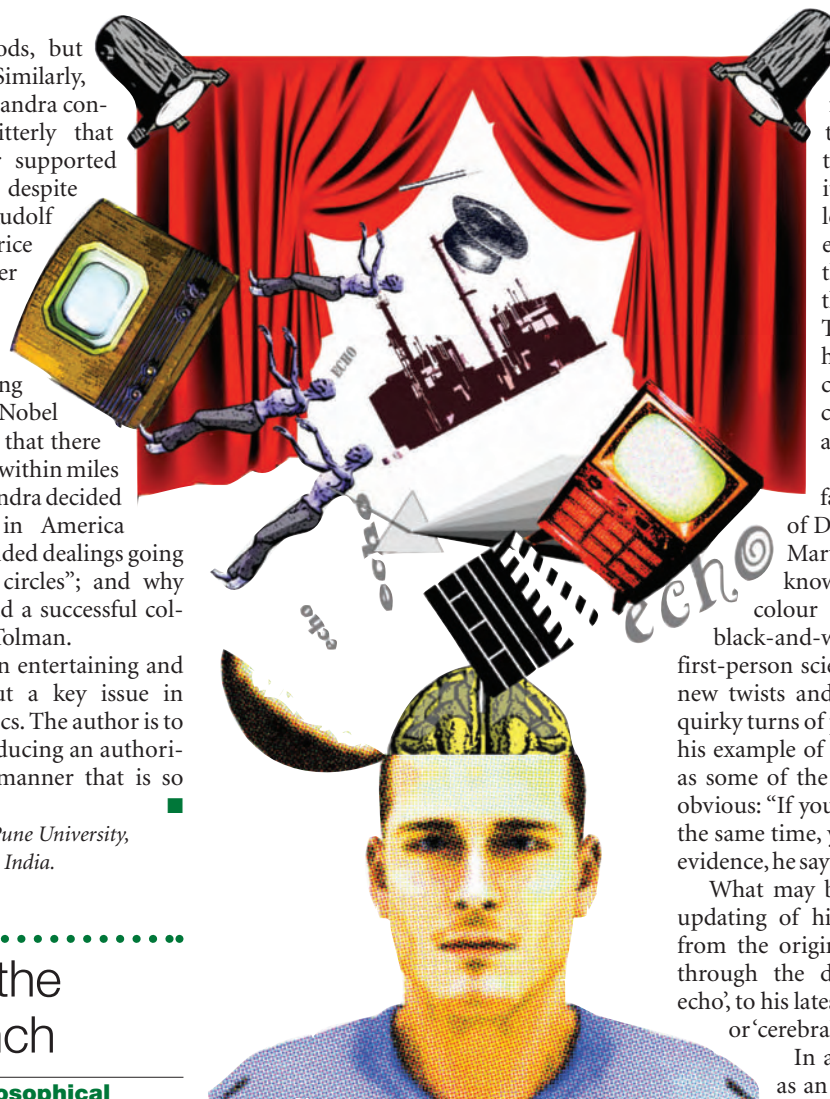
watching the magic show of conscious experiences from the safety of the cartesian theatre. Dennett contrasts those theorists to whom it is obvious that a theory that leaves out the subject cannot explain consciousness, with those to whom it is obvious that the subject has to vanish. The first type must be wrong, he says: "A good theory of consciousness should make a conscious mind look like an abandoned factory."

All these ideas will be familiar to lovers and haters of Dennett, as will his attacks on Mary the colour-scientist (who knows all there is to know about colour perception but lives in a black-and-white world) and the idea of first-person science. But he does add some new twists and, as ever, delights with his quirky turns of phrase. I laughed out loud at his example of a folk theorem as ludicrous as some of the ones that people claim are obvious: "If you burp, sneeze, and fart all at the same time, you die." But let's have some evidence, he says.

What may be less familiar is Dennett's updating of his theory of consciousness, from the original 'multiple drafts' theory, through the delightfully named 'fantasy echo', to his latest ideas of 'fame in the brain' or 'cerebral celebrity'.

In a paper originally published as an overview for a special 2001 issue of *Cognition*, he joins in the "gathering consensus" that favours 'global workspace' theory. Originally proposed by Bernard Baars in *A Cognitive Theory of Consciousness* (Cambridge University Press, 1988), this theory is explicitly based on a theatre metaphor in which the contents of consciousness are illuminated by a spotlight of attention, shining on the stage of working memory. The idea is that many specialist brain systems contribute information to the global workspace; its contents are then broadcast to the rest of the system and this global availability is experienced as a conscious state.

Doesn't this go against everything Dennett has been fighting for? He tries to explain why not. Consciousness, like fame, is not an intrinsic property of brain processes but is more like political influence or clout; conscious events are the ones that have widespread effects in the brain. So we must not think that becoming famous in the brain magically ignites the glow of conscious qualia or lets pictures into the cartesian theatre to be watched by the conscious subject; the effects are enough. Fame in the brain does not lead to consciousness — it is consciousness.



Many global-workspace theorists don't see it this way, and Dennett says they are "surrendering just when victory is at hand". The critical point hinges on what he calls the 'hard question': "And then what happens?". (This is not to be confused with Chalmers' 'hard problem', which Dennett thinks is illusory.) If you think that something has to happen next, such as entering consciousness or becoming conscious, then you are still wallowing in mystery.

But if the devil had stuck with his original multiple-drafts theory, he would surely have gone further. One of its most startling claims,

made by Dennett in *Consciousness Explained* (Little, Brown, 1991), is that “there are no fixed facts about consciousness independent of particular probes”. I took this to mean something like this: at any time there are multiple versions of representations in the brain, none of which is conscious or unconscious. Only when the system is probed — for example by asking a question or requiring an action — does one of the drafts have consequences that cause us to say, after the fact, that it was conscious. In other words, there are no conscious and unconscious streams; nothing ever ‘enters consciousness’; there

is no "crucial finish line or boundary somewhere in the brain, marking a place where the order of arrival equals the order of 'presentation' in experience".

Yet the global-workspace theorists believe in conscious and unconscious streams, in contents becoming conscious, and in looking for the neural correlates of the 'movie in the brain'. I think they still need the old devil.

Susan Blackmore is a freelance writer, lecturer and broadcaster, and a visiting lecturer at the University of the West of England, Bristol, UK.

Making progress?

FAB: The Coming Revolution on Your Desktop – From Personal Computers to Personal Fabrication

Basic Books: 2005. 288 pp. \$26, £19.99

Don Ihde

What happens to the notion of 'revolution' when it occurs every day, month or year? I suspect it comes to mean the same as 'normal' — or it becomes irritating, as constant changes end up preventing planned action.

The rhetorical cast of *FAB* is one of technofantasy and wild projection, which strangely recalls some equally extravagant medieval theological puzzles. "Can God create a stone that is too heavy for him to lift?" is paralleled here by another puzzle: can we humans build a machine that could then make a machine better than itself? This technofantasy opens *FAB*, which stands for personal fabrication: a programmable personal fabricator will be able "to make anything, including itself, by assembling its atoms. It will be a self-reproducing machine."

But, at least the author, Neil Gershenfeld — former director of the Media Lab at Massachusetts Institute of Technology (MIT) and now director of its Center for Bits and Atoms — recognizes this as a “long-standing science-fiction staple”. Boundless enthusiasm and boys-and-toys utopianism seem to be the tradition, and I had a hard time overcoming my own scepticism in order to get to a more sober assessment of what is happening.

The book arose from a popular course at MIT, “How To Make (Almost) Anything”, taught by Gershenfeld. It draws many students from diverse backgrounds, all enthusiastic to build, or personally fabricate, things. From the second chapter on, we get descriptions of some of these objects. *ScreamBody* is a sort of fuzzy soundproofed knapsack into which the user screams, but passers-by cannot hear; it contains a recording device that can then later be played and, with the bag open, be heard by others. Another is a web browser designed for use by parrots,

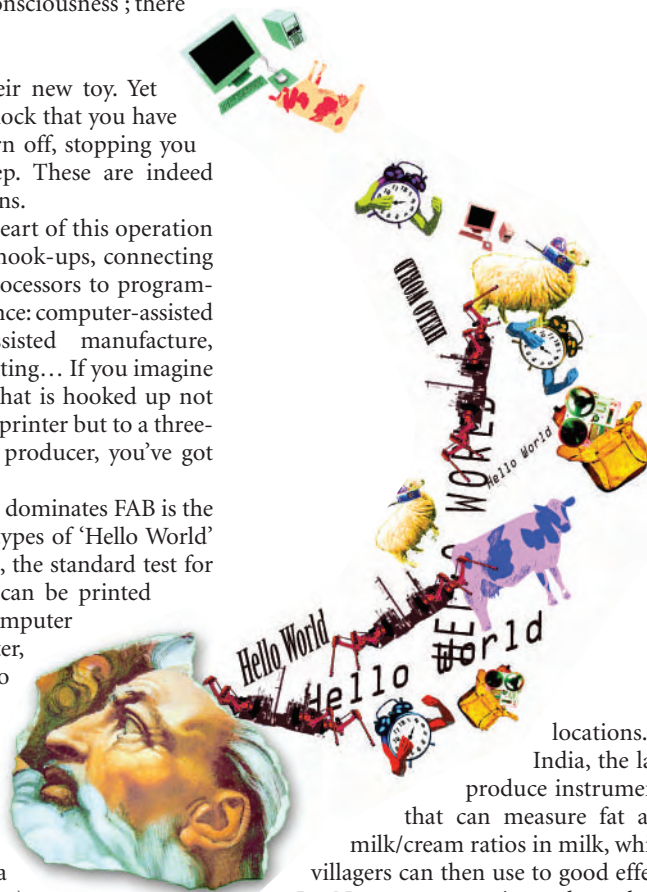
who seem to like their new toy. Yet another is an alarm clock that you have to wrestle with to turn off, stopping you slipping back to sleep. These are indeed imaginative fabrications.

It seems that the heart of this operation is a network of new hook-ups, connecting different tools and processors to programmed computer assistance: computer-assisted design, computer-assisted manufacture, computer-assisted cutting... If you imagine a desktop computer that is hooked up not to a two-dimensional printer but to a three-dimensional product producer, you've got the idea.

The trajectory that dominates FAB is the production of many types of ‘Hello World’ artefacts. Hello World, the standard test for computer languages, can be printed out by any desktop computer hooked up to a printer, but when hooked up to engraving devices it can be printed as a bas-relief, cut out as independent letters (hooked up to a laser cutter) or even in heavy material (using a supersonic water cutter).

But you could use the same water cutter to produce polycarbonate plastic bicycle frames of any design, as they do on the MIT course. The other bicycle parts are off the shelf, and when assembled the various designs could then be seen on the streets of Cambridge. I can see why this class is fun, and as a bricolage inventor myself from my Kansas childhood — my brother and I welded together a bale loader from our father's junk pile — I can begin to catch the infectious enthusiasm that permeates FAB. Yet I can't shake the doubts: what is the revolutionary future of using a US\$100,000 water cutter to make polycarbonate bicycle frames? At \$20,000 for Fab Labs exported to the Third World, the process seems rather expensive.

What is infectious, and possibly of some revolutionary importance, is the way in which Fab Labs have been exported and prepared for use by children and distant, Third World



locations. In India, the labs produce instruments

that can measure fat and milk/cream ratios in milk, which villagers can then use to good effect.

In Norway, transmitters have been developed to help herders locate their sheep, and they have been transformed into local networks for communication as well. Children from a slum learned to program their own lab, on a computer placed in a hole in the wall, often with innovations never dreamed of by the designers.

It strikes me that the ‘revolution’ that emerges is that of self-learned skills through playing with Fab Labs. Desktops hooked up to various new connections, flexible enough for eight-year-olds to invent things, in effect democratize this technical development. Such technological multistability, of course, remains unpredictable. And so much often begins in play. That is the possibility created by Fab Labs. “*Imagination au pouvoir*” was, after all, the slogan of the revolutionary Days of May in 1968 Paris. ■

*Don Ihde is in the Department of Philosophy,
Stony Brook University, New York 11794, USA.*

Climbing the evolutionary tree

The Ape in the Tree: An Intellectual and Natural History of *Proconsul*

by Alan Walker & Pat Shipman
Harvard University Press: 2005. 312 pp.
\$26.95, £17.95, €24.90

Peter Andrews

The ape in the tree is *Proconsul*, a fossil ape known exclusively from East Africa. It has many things to recommend it as a subject for a book: it is the earliest known ape in the fossil record (dating from 18–20 million years ago); it is represented by the most complete remains of any fossil ape (several partial skeletons); and, since the first remains were found some 75 years ago, it has been studied by many people of different levels of eminence. It is therefore a good subject on which to build any number of scientific edifices, whether historic, taxonomic, ecologic, phylogenetic or to do with functional morphology.

What's more, it has been at the centre of feuds, deaths, disputes over ownership, and rivalries between some of the biggest personalities in anthropology. In many of these, Alan Walker has been involved at first hand, and he brings a strongly personalized view of these events to this book. It is first and foremost an account of his own involvement with the discovery and description of some of the best specimens of *Proconsul* yet found. The account is actually written by Alan's wife, Pat Shipman, but it is curiously written in the first person singular as if by Walker, and we learn as much about him in this book as we do about the place of *Proconsul* in the evolution of the apes.

One of Walker's many gifts is his ability to illuminate subjects he is interested in and to bring new ideas to bear. His stated aim in this book is to provoke curiosity about apes and the part they have played in our evolutionary past. He recognizes that we should not view fossil apes in some sense as 'failed apes'. There is a tendency among some workers, both past and present, to try and identify fossil apes with their living survivors — chimpanzees, gorillas and orangutans — so the common ancestor of humans and chimpanzees, for example, is seen by some as having been chimpanzee-like.

Walker does not address this particular issue directly, for he is dealing with a fossil ape that lived close to the time when monkeys and apes diverged (about 20 million years ago), but he states unequivocally that the common ancestor of monkeys and apes was neither monkey nor ape but a being in itself that had its own suite of characters

and its own way of life. This enlightened approach, free of the straitjacket of descent species, is the most successful aspect of this book in documenting the way of life of *Proconsul* in all its aspects.

Having said this, Walker makes it clear right from the start that this is no scientific treatise, but a semi-popular account of early ape evolution. His mentor was John Napier, a larger-than-life character who had enormous influence on primate studies in the middle of the twentieth century. One of Napier's major works was a monograph — *The Fore-limb Skeleton and Associated Remains of Proconsul africanus* (British Museum (Natural History), 1959) — on the partial skeleton of *Proconsul* from Rusinga Island, which provides a detailed morphological description with the emphasis on the function of the animal as a whole. Napier's approach to functional morphology was hugely influential and made a great impression on Walker, who has carried on and extended this approach with the help of a succession of bright, enthusiastic students.

The study of functional morphology is based on interdisciplinary collaboration: morphology is interpreted in terms of function, function in terms of behaviour, and behaviour in terms of interaction with the environment. This requires a high level of collaboration that is entirely in line with Walker's stated intention to "give as many opportunities to up and coming young

people as I could". He is an inspired and inspiring teacher, but it is my regret, and the subject's loss, that he has not himself produced a monographic treatment of *Proconsul* and its place in ape evolution to match Napier's. The present book is not an adequate replacement, however entertaining it may be.

There is also a failure to extend the approach of functional morphology to its third aspect: interaction with the environment. There is little more than anecdotal discussion about the environment that *Proconsul* lived in, other than saying that there was forest and trees for it to move around in. The apparent link to forest environments is important for Walker's reconstruction of *Proconsul*'s way of life, for it implies a reliance on tree-living. Primates that live in woodland rather than forests, by contrast, are more terrestrial, as the tree canopies are too open and too poor in resources to support fruit-eating mammals all year round.

Walker does not mention a wealth of published environmental evidence, much of it done in the 1970s, a period of study excised from his account of *Proconsul* studies, and some that he does mention is now thought to be wrong. For example, much of the evidence for the environment at Rusinga, based on both plants and mammal faunas, shows that the area was dominated by seasonal deciduous woodlands with patches of forest that were almost certainly ephemeral both



in time and space. The strong evidence for forest environments comes from other sites: Songhor and Koru, based on abundant mammalian faunas, and Mfangano, based on the fruits of tropical forest trees. The evidence for forest at Rusinga is based on localized outcrops of forest faunas, abundant woody climbers, which typically would have been growing locally along water courses, and leaves with apparent (not yet analysed) forest affinities. The so-called fruit-and-nut bed, which Walker mentions several times as providing evidence of forest, is actually dominated by plant species with deciduous woodland affinities. The idea of *Proconsul* as a slow climber subsisting mainly on a diet of soft fruits, as indicated by its functional morphology, is less viable if it lived in a deciduous woodland environment, rather than a forest.

Walker's semi-popular account of this important fossil ape will be an accessible and entertaining read for the educated layman. It will also be a useful guide to students learning about how research is conducted, although they need to be aware of the idiosyncracies in this highly personalized account. We can live in hope that Walker will do justice to his great abilities and will yet produce a definitive account of this ape in the tree. ■

Peter Andrews is in the Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.

The evolution of scientific thinking

Reef Madness: Charles Darwin, Alexander Agassiz, and the Meaning of Coral

by David Dobbs

Pantheon Books: 2005. 306 pp. \$25

Rachel Wood

Reefs have long fascinated natural historians and geologists for their unearthly beauty, as well as their ability to produce prodigious amounts of carbonate sediment. Yet reefs offer more than their share of paradoxes. How do coral-reef islands seemingly grow from great depths in the middle of the oceans? What controls the production of all this limestone? And why do so many reefs form necklaces strung across the Pacific? These questions troubled the minds of nineteenth-century scientists and philosophers. The long, tortured and often sad history of how the 'coral reef problem' was finally solved is laid bare in this eloquent and thoughtful book.

We are first introduced to the key figures: Louis Agassiz, his son Alexander and Charles Darwin. The book explores in detail

the meteoric rise and fall of the arrogant, narcissistic and charismatic Louis, and his relationship with his shy and diligent son. Louis, a palaeontologist, was the first to propose that an ice age could explain many features of the Earth's surface. In the mid-nineteenth-century United States, which was hungry for pioneering personalities that embodied the spirit of a young nation, Louis' enormous energy and persuasive vision enabled him to found many scientific institutions that still thrive today. Darwin, by contrast, is presented as a distant figure, more a man with a following than a personality.

It was Louis' failure to embrace the implications of Darwin's *Origin of the Species*, instead tenaciously holding on to his idealistic logic that God had created all species whole and immutable, that toppled him from his pedestal in American scientific and Boston blue-blood society. In just five years Louis was transformed from being seen as the prince of US science to an archaic reactionary.

As a young man, Alexander was caught between paternal loyalty and a great personal need for objective and careful analysis. He rejected his father's florid unscientific style but was equally unhappy with Darwin's theorizing, which was based, as Alexander saw it, on equally unacceptable flights of imagination, rather than on the tireless compilation of meticulously observed facts.

Darwin had been greatly influenced by the geologist Charles Lyell, whose major tenet was that geological structures, rather than being the result of past catastrophes, were instead formed by the constant action of slow and gradual processes observable today. It is easy to see how such thinking led to Darwin's formulation of another time-based phenomenon: the origin of multiple species through common descent by means of natural selection. Darwin's observation of Pacific reefs, together with his harrowing experiences of the Earth-moving forces of earthquakes during the voyage of the *Beagle*, led him to propose that oceanic reefs and atolls were formed on subsiding foundations, such as volcanoes. His imaginative intellect saw here another result of small changes that could account for otherwise complex patterns: that there was a dynamic relationship between the reefs and their foundations that seemed to shape them. But proof for this theory was there none.

So the stage was set for a century-long battle for the roles of empiricism, theorizing and imagination in scientific endeavour. In 1876, the Scottish oceanographer John Murray had put forward an alternative coral-

reef hypothesis. He proposed that reefs grew not on their own debris, but on the accumulation of sediment derived from plankton and other non-coral skeletons. Alexander Agassiz was immediately drawn to this idea: it was born of neither his father's idealism nor Darwin's over-simple theorizing, but seemed instead to be based firmly on observable fact. Alexander, grief-stricken by the early loss of his wife and other relatives, had finally found his *raison d'être*: to disprove Darwin's theory.

Made hugely wealthy from the ownership of copper mines, Alexander was able to embark on a massive programme of travel. For decades he made detailed observations of all the known coral reefs. He was convinced that each reef was a



unique outcome of forces that combined in different ways: no single grand theory was required. By 1902, he believed he had proved Murray's theory. Yet astonishingly, after his death in 1910, no manuscript that summarized his massive accumulation of observations was ever found. It seems that for Alexander, the chase was everything.

The truth is that there were never enough facts to go on. All the protagonists had died long before the origin of reefs was resolved by deep drilling in the 1950s. The penetration of Eniwetok atoll down to its volcanic foundations revealed some 1,500 metres of coral reef limestone, proving Darwin right.

Reef Madness is more than a narrative about the victory of empiricism, the power of observation as the evidence of truth, over the philosophy of belief. It is also a beautiful illustration of how theories can, or must, be built slowly and painfully, brick by brick, by a dynamic combination of both imaginative leaps and factual observation. It would be foolish to think that this debate has ended: as

the author points out, 150 years later, over half of Americans continue to believe that God has either created species or directs evolution in some way. It seems that science and the human condition will always need a beautiful idea. ■

Rachel Wood is at Schlumberger Cambridge Research and in the Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.

With the future in mind

The 21st-Century Brain: Explaining, Mending and Manipulating the Mind

by Steven Rose

Jonathan Cape: 2005. 352 pp. £20

Published in the US as *The Future of the Brain* (Oxford University Press, \$28)

John C. Marshall

Steven Rose is both a distinguished neuroscientist and a conscientious citizen who has long been active in left-wing causes. Reflecting this duality, *The 21st-Century Brain* is surely the only book on neuroscience in which the index entry for 'consciousness' includes both 'class' and 'feminist' as sub-headings. It is also the only one that the author says "is definitely not about offering some dramatic new 'theory of consciousness'". The omens, then, are good.

Rose offers a fairly comprehensive, albeit simple, account of our current understanding of the brain, of where research might lead us next, and of the ethical issues that already arise from the application of neuroscience and that will multiply rapidly in the near future.

He takes us on the evolutionary journey from the metabolic soup of protocells in sea water to the 100 billion richly interconnected nerve cells that make up the human brain. Yet all the while he insists that the physical brain enables, but does not determine, mental life. The location allows Rose to stress how individual life histories are "shaped by culture, society and technology", although this observation, however true, does little to bridge the gap between brain and mind.

Indeed, when the chips are down, Rose's position is indistinguishable from one form of dualism: "In the lab we can all aspire to objectivity, examining the workings of other brains — or even imagining our own — yet we go home in the evening to our subjective, autobiographical world, and aspire to make personal sense of our lives and loves." Neuroscientists, he declares, "must learn to live with this contradiction". He is even prepared to believe that the natural sciences are perhaps intrinsically incomplete and must be complemented by the kinds of knowledge we gain from the arts and humanities. Certainly, he sees no reason to believe E. O. Wilson's claim that art, literature and music arose as evolutionary strategies for attracting mates.

I have every sympathy for Rose's position here, but I wonder whether neodarwinism can survive in its present form if we are allowed to pick and choose which behaviours (or their underlying genes) have been subject to natural selection. For example,

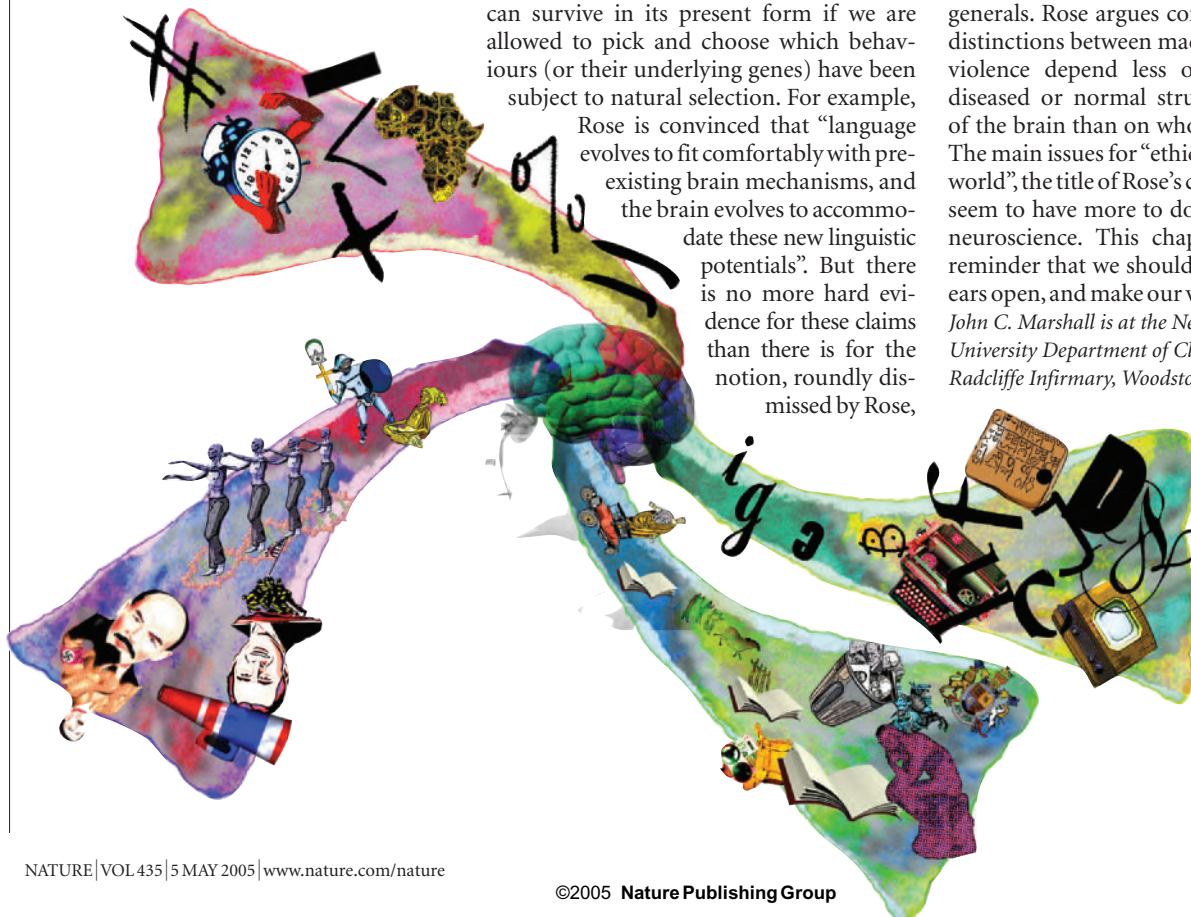
Rose is convinced that "language evolves to fit comfortably with pre-existing brain mechanisms, and the brain evolves to accommodate these new linguistic potentials". But there is no more hard evidence for these claims than there is for the notion, roundly dismissed by Rose,

that human nature was fixed in the Pleistocene 600,000 to a million years ago. At some point we may understand how the roughly 1% genetic difference between chimpanzees and humans is responsible for such a cognitive gulf between the two species, including the capacity for language; at present we are in the realm of 'just so' stories.

Even with respect to Rose's own area of research — learning and memory — the lack of detail in *The 21st-Century Brain* makes the line of argument confusing to the outsider. Rose notes that "because non-human animal brains work very much in the same way as do our own, I am able to work with experimental animals to analyse the molecular and cellular processes that occur when they learn some new skill or task". A little later, however, we discover that "different species have widely differing skills" in what they can and cannot learn. No doubt the two claims are compatible, but it would have been helpful to have the reconciliation spelled out. The section on autobiographical memory only confuses the issue further. Rose reports the position of the psychologist Endel Tulving that memories are not permanently stored in the brain, but rather "actively called into being, in the process of recall". I suspect that Rose would disagree, but sadly he fails to comment.

In the concluding chapters, Rose moves into the controversial realms of mental illness, psychopharmacology, thought control, neurogenetics and biocybernetics. These pages are appropriately sceptical of the motives of drug companies, politicians and generals. Rose argues convincingly that the distinctions between mad, bad and justified violence depend less on the supposedly diseased or normal structure and activity of the brain than on who wields the power. The main issues for "ethics in a neurocentric world", the title of Rose's concluding chapter, seem to have more to do with politics than neuroscience. This chapter is a welcome reminder that we should keep our eyes and ears open, and make our voices heard. ■

John C. Marshall is at the Neuropsychology Unit, University Department of Clinical Neurology, Radcliffe Infirmary, Woodstock Road, Oxford OX2 6HE, UK.



Literature red in tooth and claw

Madame Bovary's Ovaries: A Darwinian Look at Literature

by David P. Barash & Nanelle R. Barash
Delacorte Press: 2005. 272 pp. \$24

Michel Raymond

Book titles that feature the possessions of celebrities have recently been in vogue. We've seen *Marengo: The Myth of Napoleon's Horse* (history), *Mendel's Dwarf* (fiction) and even *Napoleon's Buttons* (chemistry). Now we've got *Madame Bovary's Ovaries*, an unusual — and unusually readable — effort at literary criticism. This book attempts to interpret fiction in terms of evolutionary biology.

This may not sound promising, but the idea is more robust than it might at first appear. In 1973 the great geneticist Theodosius Dobzhansky remarked: "Nothing in biology makes sense except in the light of evolution." These words are now a touchstone for biologists seeking to understand the surrounding world. But in the past thirty

years, his comment has acquired an unexpected breadth: biology is currently expanding, or, more precisely, the classical boundaries between biology and the social sciences are fading away.

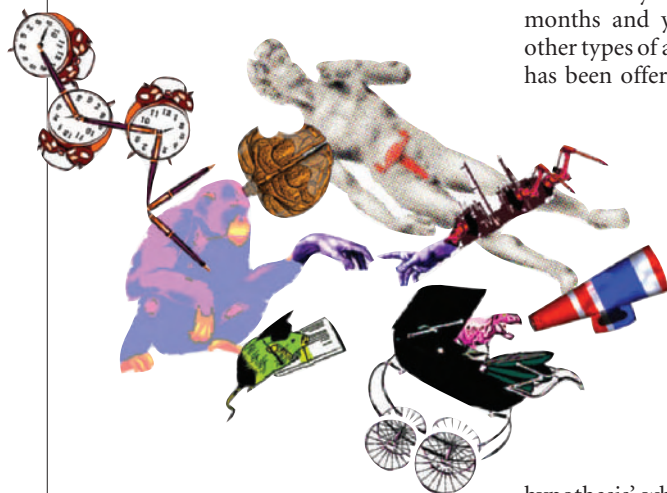
Animal behaviours and cultures have been brilliantly studied by biologists. Human culture is complex, but humans are animals too, so Dobzhansky's adage also applies to us. These days, at conferences on human behaviour, you run into evolutionary psychologists, stumble across darwinian historians and evolutionary anthropologists, listen to talks on darwinian medicine, and so on. Human sciences and the humanities have now been extirpated from the old non-historical theories, incompatible with the idea of evolution, that prevailed during the twentieth century. Goodbye Freud, Piaget, Levi-Strauss, Eliade, Chomsky and company.

Literature is a cultural product from a living species, so Dobzhansky's words apply, and this book is best understood within the context of a revival of the humanities. But in the evolutionary study of literature, there are two sides to consider: the writer and the reader. Unfortunately, nothing in *Madame Bovary's Ovaries* addresses the writer's motivations: why do people spend hours, days, months and years creating literature and other types of art? One intriguing possibility has been offered: Geoffrey Miller's 'display

Misérables and Harry Potter), and male-male competition for access to higher social status (*Huckleberry Finn*).

As a result, *Madame Bovary's Ovaries* explores various aspects of human mating strategies, rooting human behaviours within the animal repertoire. But where most such books rely on scholarly papers and monographs to ground the various points within a shared and robust scientific knowledge, these authors use literature sources as references. And why not? After all, a successful novel should contain important social information. And if this novel is still widely read a century later or in various cultural groups, its social information probably concerns a universal human topic, or at least one found in most cultures (such as male-male competition for higher social status).

Unfortunately, the various genres of literature have not been commented on in this book. An evolutionary book could have provided some welcome insights on, for example, the historical origin of the novel in various human cultures, and its unclear relationships with other kinds of literature,



such as myths or fairy tales. Perhaps this could be the subject of another book?

Madame Bovary's Ovaries lies at the crossroads between literary studies and biology, and has much to offer students of either subject. A biologist reader will be surprised at his expanded field of action, and unless he has already done so, he must now catch up with some basic literature; the good news is that he has in his hands a good guide to choosing a novel according to his own marital, familial or social concerns. A literary reader will be given a tour of some pivotal concepts of evolutionary biology, and learn how to use them to explore more new fields.

There is a profusion of books devoted to literature. This one does not talk about why books are written, but about the (darwinian) motives to read novels and other kinds of literature. We don't really know why the authors wrote this book. However, when reading it, it is easy to feel that it provides an interesting addition to our knowledge of human culture. ■

Michel Raymond is at the Institute of Evolutionary Sciences, University of Montpellier II, Montpellier 34095, France.

hypothesis', which states that people tend to produce art as a strategy for broadcasting courtship displays to a wide variety of potential partners.

It is easier to explain why people spend so much time reading literature. If someone sits for a long time with a book in his hands, he is probably extracting something interesting from it. In evolutionary currency, 'something interesting' relates to reproduction, either directly or indirectly (social competition, for example). Accordingly, the authors of *Madame Bovary's Ovaries* place literature into several darwinian drawers, such as male sexual jealousy (*Othello*), how to find the best mate for a female (the novels of Jane Austen), adultery (*Madame Bovary*), the reduced parental investment when parent-child relatedness is low (*Cinderella*, *Les*



Ghost of speciation past

Thomas D. Kocher

Lurking in the rivers of Botswana are the remnants of a diverse flock of cichlid fishes, whose origins can be traced to a lake that vanished more than 2,000 years ago.

Just add water. A simple recipe, but when applied to cichlid fishes it reliably causes spectacular evolutionary radiations of hundreds of new species. Generalist riverine cichlids have given rise to flocks of highly specialized fishes in each of the major lakes in East Africa. Now, on page 90 of this issue, Joyce and colleagues¹ describe a previously unrecognized radiation of cichlids in the extinct Lake Makgadikgadi, whose descendants still haunt the rivers of southern Africa today.

Joyce and colleagues' evidence¹ consists of mitochondrial DNA sequences determined from fishes collected from rivers throughout southern Africa. Rivers in East Africa typically hold just one haplochromine cichlid species (haplochromine is the name given to one of the major lineages of cichlid fishes in Africa). But the Okavango, Cunene, Limpopo and upper sections of the Congo and Zambezi rivers in southern Africa each harbour several species of haplochromines. Evolutionary analyses show that these fishes share a common ancestor quite separate from that which gave rise to the better known radiations in lakes Malawi and Victoria (Fig. 1). The diversity of these riverine cichlids peaks in the Okavango delta, which geological evidence suggests was, until about 2,000 years ago, part of a lake larger than Switzerland.

The extent of this newly recognized adaptive radiation is truly impressive. The remnants of the Lake Makgadikgadi flock display a morphological diversity nearly as great as that of the extant cichlids in lakes Malawi and Victoria. The flock includes ambush predators with long heads and jaws, and snail crushers with deep heads and massive teeth. Only a few specialized morphologies are missing — such as those of pelagic plankton eaters and benthic algal scrapers — probably because these niches are not available in the remaining river habitat.

What is it about lakes that facilitates these rapid radiations? Certainly, large lakes provide a range of new ecological habitats not present in rivers. The still waters of lakes probably also reduce rates of gene flow between different groups of fish, enabling them to diverge genetically from each other and eventually to speciate. But other groups of fishes have not responded to these physical factors with similar bursts of speciation. It seems there is also something special about haplochromine cichlids that causes them to

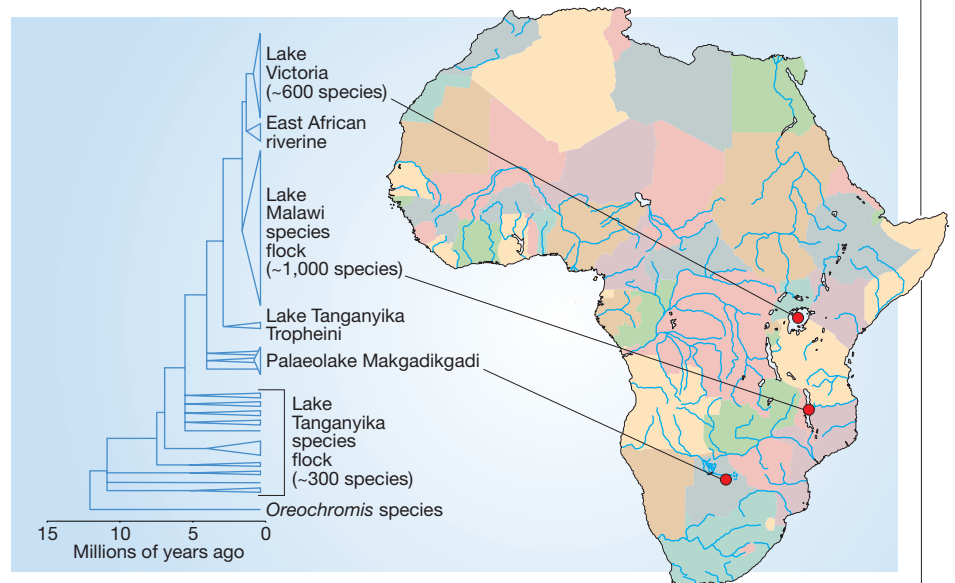


Figure 1 Repeated diversification of cichlid fishes in southeast Africa. Mitochondrial DNA sequences reveal a radiation of cichlids in Lake Tanganyika beginning 7–9 million years ago. Over the past 3–4 million years, one of these lineages has spread through the rivers of southern and eastern Africa. These haplochromine cichlids gave rise to independent radiations in lakes Victoria and Malawi, and, as Joyce *et al.*¹ now describe, in the palaeolake Makgadikgadi. In Lake Tanganyika, these fishes also gave rise to a recent radiation of the Tropheini. The tree is rooted with tilapia (genus *Oreochromis*), a popular food fish that is closely related to the haplochromine cichlids of the region. (Tree modified from ref. 10.)

diversify in these large lakes. What are the key innovations that make them particularly prone to rapid adaptive radiation?

A long-standing hypothesis is that cichlids have a malleable feeding apparatus that can be quickly altered to take advantage of new feeding opportunities². Cichlids have a second set of jaws in the back of the throat — the pharyngeal jaws — that process food before it enters the gut. The oral jaws are therefore relatively free to evolve specializations for acquiring food. This decoupling of functions has clearly contributed to the rapid evolution of feeding specializations in each of the species flocks that inhabit lakes.

Another key innovation is maternal mouth-brooding. Most non-haplochromine cichlids lay large numbers of small eggs on the lake or river bottom, and both parents cooperate in the care of the eggs and fry. In contrast, female haplochromines lay small numbers of large eggs, and carry the eggs in their mouths for several weeks until the young are released to fend for themselves. Haplochromine males provide no post-spawning care of their offspring.

This unequal parental investment results in strong sexual selection. Males must compete for females, and as a consequence have evolved an astonishing variety of nuptial colorations and behaviours to attract mates. These differences in colour are often enough to prevent interbreeding (hybridization) among existing species³. But it is not clear whether sexual selection on male traits is sufficient to drive speciation in the first place. Other genetic conflicts, including those between females and their offspring over the sex ratio, could also be important⁴.

Joyce and colleagues¹ suggest that hybridization may itself have played a part in the origin of the Lake Makgadikgadi species flock. This lake was probably colonized by distantly related cichlids from previously unconnected rivers. If these lineages hybridized, the resulting pool of genetic diversity might have facilitated the generation of new morphological or behavioural characteristics, enabling the colonization of newly available ecological niches in the lake⁵. This hypothesis is difficult to prove, but may be testable through analysis of variation

in the genes underlying important adaptive traits.

The present situation in the Okavango delta may also provide some context for discussing the origin of today's Lake Victoria flock. Although some geological data suggest that the Lake Victoria basin was completely dry 15,000 years ago, molecular evolutionary studies generally suggest that the flock is much older (100,000 years)^{6,7}. As with Makgadikgadi, it seems likely that several genetically and morphologically diverse cichlids survived the drying of Lake Victoria, and have been able to re-radiate into more than 500 species during the 15,000 years since the lake refilled⁸.

The high rates of speciation observed in these African cichlids are almost beyond belief⁹, but the evidence is clear. The discovery of yet another species flock¹ emphasizes the importance of the phenomenon, and reinforces the utility of these fishes for

studying evolutionary mechanisms. African cichlid fishes represent about 5% of all vertebrate species. A synthesis of the mechanisms responsible for their spectacular radiation is essential if we are to fully appreciate the origins of vertebrate diversity. ■

Thomas D. Kocher is at the Hubbard Center for Genome Studies, University of New Hampshire, Durham, New Hampshire 03824, USA.

e-mail: tom.kocher@unh.edu

1. Joyce, D. A. *et al.* *Nature* **435**, 90–95 (2005).
2. Liem, K. *Syst. Zool.* **22**, 425–441 (1974).
3. Knight, M. E. & Turner, G. F. *Proc. R. Soc. Lond. B* **271**, 675–680 (2004).
4. Kocher, T. D. *Nature Rev. Genet.* **5**, 288–298 (2004).
5. Seehausen, O. *Trends Ecol. Evol.* **19**, 198–207 (2004).
6. Seehausen, O. *et al.* *Proc. R. Soc. Lond. B* **270**, 129–137 (2003).
7. Verheyen, E., Salzburger, W., Snoeks, J. & Meyer, A. *Science* **300**, 325–329 (2003).
8. Seehausen, O. *Proc. R. Soc. Lond. B* **269**, 491–497 (2001).
9. Coyne, J. A. & Orr, H. A. *Speciation* (Sinauer, Sunderland, MA, 2004).
10. Salzburger, W. & Meyer, A. *Naturwissenschaften* **91**, 277–290 (2004).

Developmental biology

Morphogens hitch a greasy ride

Richard S. Mann and Joaquim Culi

Morphogen proteins guide the development of many tissues in animals, but how are these insoluble proteins ferried around the body? A well-known group of lipid transporters might be the answer.

Every once in a while, a scientific discovery results in the marriage of two previously disparate fields. On page 58 of this issue, such a romance is suggested by Eaton and colleagues¹. On the one hand, there are morphogens — signalling proteins that play an essential part in animal development by inducing specific cellular fates in a concentration-dependent manner^{2,3}. How these proteins move from the cells that synthesize them to neighbouring cells to create a concentration gradient is an area of intense debate. On the other hand, there are lipoprotein particles, most famous for their role in lipid transport and heart disease. The marriage, witnessed by intriguing experiments described in the new paper, raises the possibility that lipoprotein particles carry morphogens as they move through tissues during development.

Although biologists have had much success in measuring the biological effects of morphogens, understanding how these proteins spread through tissues to generate a concentration gradient has proved more challenging. This is in part because many morphogens are highly insoluble, which hampers their diffusion through the hydrophilic environment of tissues. In particular, two morphogens, the Hedgehog (Hh) and Wnt/Wingless (Wg) proteins, are attached to palmitic acid, and Hh is also covalently linked to cholesterol⁴; these lipids

would promote the association of the two proteins with cell membranes.

Several models have been proposed to explain how morphogens move through tissues. Some suggest that they are actively transported by cells — through long cellular extensions called cytonemes, for example, or by being passed from cell to cell through cycles of secretion and uptake^{2,3}. Although these models satisfactorily deal with morphogen insolubility, it has been difficult to confirm their role in morphogen spreading and activity. In contrast, accumulating biological evidence and mathematical modelling favour the idea that morphogens diffuse passively along the surface of cells, aided by local interactions with other molecules^{2,3}. For such 'restricted diffusion' to be possible, lipid-modified morphogens must somehow acquire a relatively soluble form.

One potential solution, suggested previously by the Eaton group⁵, posits that morphogens move in membranous vesicles called argosomes that are derived from morphogen-producing cells. Now the Eaton and Thiele labs¹ have revised this model by suggesting that morphogens move via large particles better known for their role in transporting lipids — the lipoproteins (Fig. 1). If they are right, this connection would contribute significantly to our understanding of morphogen spreading, and suggests a new role for lipoproteins during animal development.

Lipoproteins are large, globular macromolecular complexes composed of a central core of lipids surrounded by an outer layer of polar phospholipids, cholesterol and specialized proteins called apolipoproteins⁶. Most lipoproteins are synthesized in the liver and intestine of mammals and in the fat-body of insects, from where they move to peripheral tissues through the blood and lymph (the haemolymph in insects) to regulate lipid levels throughout the body.

Using their lipid adducts as anchors, Wnt and Hh proteins could in principle easily reside in the outer phospholipid layer of lipoproteins. Indeed, using biochemical fractionation of fruitfly (*Drosophila*) tissues, Eaton and colleagues¹ show that, although most Wg and Hh proteins co-fractionate with cellular membranes, a significant percentage partitions into the lipoprotein fraction. Moreover, lipoproteins can be co-immunoprecipitated with Wg and Hh, suggesting a tight association.

So, lipoproteins and morphogens have been seen together, but is this relationship a meaningful one? Eaton and colleagues suggest that it is, because lipoproteins seem to be required for the accurate establishment of morphogen gradients and activity in *Drosophila* larvae. The authors used a process called RNA interference (RNAi) to decrease apolipoprotein synthesis in the *Drosophila* fat-body — thus decreasing lipoprotein concentrations throughout the animal. This reduced the range of action of Hh and Wg in the imaginal discs, where morphogen activity is best characterized. Moreover, Hh accumulated at higher than normal levels in cells near the Hh source, perhaps as a consequence of reduced diffusion.

Although these findings are intriguing, the systemic way in which lipoproteins are synthesized and circulated, combined with their role in lipid metabolism, makes it difficult to design a watertight experiment. The authors describe two controls that help strengthen their conclusions. First, incubating the RNAi-treated discs with purified lipoproteins at least partially restored the long-range signalling activity of Hh. Second, growing larvae on a reduced-lipid diet did not affect morphogen activity or diffusion. Both experiments support the idea that the RNAi-induced defects are due to lower numbers of lipoprotein particles and not to a secondary effect such as reduced lipid levels.

Beyond providing a new mechanism by which lipophilic morphogens might spread through tissues (Fig. 1), the marriage between lipoprotein metabolism and animal development, if correct, hints at additional relationships. For example, the diffusion of most morphogens also requires heparan sulphate proteoglycans (HSPGs) — large molecules consisting of protein and carbohydrate that are found on the cell surface⁷. Lipoproteins are already known to interact

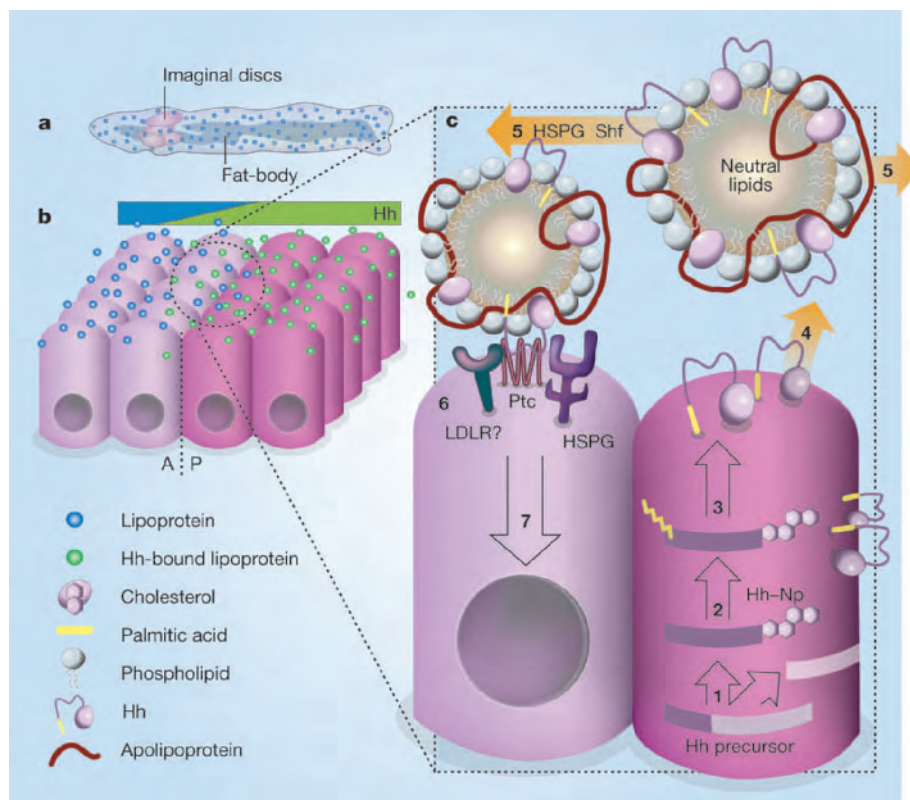


Figure 1 Morphogen movement. How a gradient of Hedgehog (Hh) protein might form, as proposed by Eaton and colleagues¹. **a**, Lipoprotein particles are made in the fat-body and distributed throughout a fruitfly larva, coming into contact with the imaginal discs. **b**, A blow-up of an imaginal disc close to its anterior (A)–posterior (P) boundary. Posterior cells secrete Hh, producing a gradient of Hh-charged lipoprotein particles and a counter gradient of uncharged particles. **c**, Blow-up of A and P cells, showing some of the steps that produce and transduce the Hh gradient. 1: The Hh precursor cleaves itself and is added to cholesterol to form Hh–Np. 2: Palmitic acid is added. 3: The protein is transported to the plasma membrane. 4: The Discharged protein might help to load Hh into lipoproteins. 5: Loaded particles diffuse somewhat, with the help of heparan sulphate proteoglycans (HSPGs) and the Shifted (Shf) protein^{19,20}. 6: Charged lipoproteins bind to receiving cells, aided by the Hh receptor Patched (Ptc) and possibly HSPGs and low-density-lipoprotein receptors (LDLRs). 7: The signal is transmitted to the nucleus.

with HSPGs, which are also important in lipid metabolism⁸. Thus, HSPGs might affect morphogen diffusion, at least in part, by interacting with their lipoprotein carrier. Consistent with this idea, diffusion of a form of Hh that cannot be modified by lipids is independent of HSPGs⁷.

Another fascinating potential link concerns how target cells might bind and take up morphogen–lipoprotein complexes. Lipoproteins have their own receptors, the low-density-lipoprotein receptors (LDLRs), and certain members of this protein family are essential for receiving the Wg signal^{9,10}. But it is not clear whether they interact directly with Wg¹¹. Perhaps instead they bind the lipoprotein moiety of Wg–lipoprotein particles. Other LDLRs might similarly contribute to morphogen activity^{12,13}.

Yet another exciting possibility concerns the mechanism that delivers Hh and Wg from the membranes of the cells that produce them to lipoproteins. Perhaps morphogens associate reversibly with lipoproteins, thus generating a source of morphogen-charged

particles close to morphogen-producing cells. Alternatively, a more active mechanism may exist. Consistent with this notion, the protein Dispatched — a member of the sterol-sensing receptor family — is essential for the release of cholesterol-modified Hh from Hh-producing cells, but is not required for the release of a non-cholesterol-modified form¹⁴. Thus, Dispatched might be involved in charging lipoproteins with Hh. A similar mechanism involving lipid rafts — specialized regions of the cell membrane — has also been suggested for Wg secretion¹⁵.

Despite the many possibilities suggested by this model, more questions are raised, in part because both fields have been so productive. How efficiently are peripheral tissues, such as the lumens of *Drosophila* imaginal discs, bathed by lipoproteins? Are lipoproteins the sole carriers for Hh and Wg, or are other mechanisms, such as the formation of micelle-like Hh multimers¹⁶, also involved? Because Wg and Hh sometimes act on the same cells, do individual lipoprotein particles carry both morphogens at the same



100 YEARS AGO

On Sunday, the President of the French Republic entertained the King at the Elysée at a dinner party, at which 120 guests were present. The guests included distinguished authors, artists, musicians, and other representatives of intellectual activity, almost exclusively members of the Institute of France. By inviting leaders of literature, art, and science to meet the King, graceful recognition was given of the high place occupied by the muses in the polity of the Republic. In the days when sheer muscular force was the mainstay of a nation, bodily strength and prowess were rightly regarded as recommendations for Court favours; but now that brain-power instead of muscle determines the rate of national progress, the State that desires to advance must foster all the intellectual forces it possesses. This principle is well understood in France, and is also clearly recognised in Germany, where every man who makes notable contributions to knowledge of any kind, assists industrial progress, or creates works of distinguished merit, whatever they may be, is sure to receive personal encouragement from the Emperor. The presence of these leaders of thought is a striking characteristic of the German Court; while, on the other hand, their absence, and the overpowering influence of military interests, are distinguishing features of Russian, and, let us add, of British Court functions.

ALSO:

Satisfactory progress and general prosperity form the key-note of the report of the Zoological Gardens at Giza for the past year. The report is illustrated by the reproduction of a most interesting photograph of an aardvark, or ant-bear, slightly marred by the effect of a shadow by the side of the nose. From *Nature* 4 May 1905.

50 YEARS AGO

“Mathematical Association Annual Meeting.”

The first item in the afternoon session was a discussion of “The Disadvantages of a Mathematical Education”, led by Mr. W. O. Storer (Department of Education, University of Birmingham); Mr Storer thought that the logical training supposedly given by mathematics might be arid, and that mathematical insistence on accuracy might lead to intellectual arrogance. Some members were reluctant to accept these inferences.

From *Nature* 7 May 1955.

time? Also, because Eaton and colleagues studied only *Drosophila* larvae, and examined only a few target genes, it is not known whether this mechanism also operates in *Drosophila* and vertebrate embryos. Suggestively, however, insect egg yolk contains abundant lipoprotein, and lipoprotein synthesis occurs in the yolk-sac visceral endoderm of early mouse embryos^{17,18}.

We also wonder which of the several varieties of mammalian lipoproteins could be involved in morphogen diffusion, and if they provide additional specificity. And, finally, could the diffusion of morphogens not associated with lipids, such as members of the transforming growth factor- β /activin/Dpp family, also be mediated by lipoproteins? Thus, from this marriage, we expect many interesting progeny. ■

Richard S. Mann is in the Department of Biochemistry and Molecular Biophysics, Columbia University, 701 West 168th Street, HHSC 1104, New York, New York 10032, USA.
e-mail: rsm10@columbia.edu

Joaquim Culi is at the Centro de Biología Molecular 'Severo Ochoa', Universidad

Autónoma de Madrid, 28049 Madrid, Spain.

e-mail: jculi@cbm.uam.es

1. Panáková, D., Sprong, H., Marois, E., Thiele, C. & Eaton, S. *Nature* **435**, 58–65 (2005).
2. Zhu, A. J. & Scott, M. P. *Genes Dev.* **18**, 2985–2997 (2004).
3. Tabata, T. & Takei, Y. *Development* **131**, 703–712 (2004).
4. Mann, R. K. & Beachy, P. A. *Annu. Rev. Biochem.* **73**, 891–923 (2004).
5. Greco, V., Hannus, M. & Eaton, S. *Cell* **106**, 633–645 (2001).
6. Jonas, A. in *Biochemistry of Lipids, Lipoproteins and Membranes* 4th edn (eds Vance, D. E. & Vance, J. E.) 483–504 (Elsevier, Paris, 2002).
7. Lin, X. *Development* **131**, 6009–6021 (2004).
8. Kolset, S. O. & Salmivirta, M. *Cell Mol. Life Sci.* **56**, 857–870 (1999).
9. Hussain, M. M., Strickland, D. K. & Bakillah, A. *Annu. Rev. Nutr.* **19**, 141–172 (1999).
10. He, X., Semenov, M., Tamai, K. & Zeng, X. *Development* **131**, 1663–1677 (2004).
11. Wu, C. H. & Nusse, R. *J. Biol. Chem.* **277**, 41762–41769 (2002).
12. Spoelgen, R. *et al. Development* **132**, 405–414 (2005).
13. McCarthy, R. A., Barth, J. L., Chintalapudi, M. R., Knaak, C. & Argraves, W. S. *J. Biol. Chem.* **277**, 25660–25667 (2002).
14. Burke, R. *et al. Cell* **99**, 803–815 (1999).
15. Zhai, L. *et al. J. Biol. Chem.* **279**, 33220–33227 (2004).
16. Zeng, X. *et al. Nature* **411**, 716–720 (2001).
17. Kawooya, J. K., Osir, E. O. & Law, J. H. *J. Biol. Chem.* **263**, 8740–8747 (1988).
18. Farese, R. V. Jr *et al. J. Lipid Res.* **37**, 347–360 (1996).
19. Glise, B. *et al. Dev. Cell* **8**, 255–266 (2005).
20. Gorfinkiel, N., Sierra, J., Callejo, A., Ibanez, C. & Guerrero, I. *Dev. Cell* **8**, 241–253 (2005).

Planetary science

Saturn's retrograde renegade

J. Brad Dalton

Data from the Cassini–Huygens mission provide convincing evidence that the saturnian moon Phoebe formed elsewhere in the Solar System, and was only later captured by Saturn's gravitational pull.

In the next few years, the Cassini–Huygens spacecraft will return a wealth of data from its orbital tour of Saturn's rings and moons. Although many images are released to the public almost immediately, it takes time to conduct detailed scientific analyses of the observations. Stunning results from early encounters with the moons are now beginning to trickle in.

On 11 June 2004, shortly before it permanently entered Saturn's orbit, Cassini–Huygens made its closest approach to the planet's outermost moon, Phoebe (Fig. 1; see also Box 1, overleaf). Phoebe is a small (around 220-km-diameter), irregular satellite in an eccentric, inclined, retrograde orbit — moving counter to the direction of the planet's rotation, in contrast to the more common, prograde, movement. Scientists have long suspected^{1,2} that Phoebe did not form alongside Saturn and its regular satellites, which occupy roughly circular, low-inclination, prograde orbits. Two papers in this issue^{3,4} provide additional evidence for this hypothesis.

Johnson and Lunine³ (page 69) use density and volume calculations, determined from navigation and imaging systems on board Cassini, to suggest that Phoebe is more similar to objects found in the Kuiper belt,

a collection of small, icy bodies beyond Neptune, than it is to its siblings in orbit about Saturn. Clark *et al.*⁴ (page 66) interpret spectral observations from Cassini's Visual and Infrared Mapping Spectrometer to reveal a complex array of surface materials ranging from minerals to volatile organic compounds.

A number of these compounds have not

yet been observed on the regular saturnian satellites, but are typically found in primitive bodies of the outer Solar System, such as comets, certain types of asteroid and, again, objects in the Kuiper belt. As the Kuiper belt is thought to contain material left over from the formation of Saturn, Uranus and Neptune, such an origin for Phoebe still keeps it in the family, albeit as a more distant, 'prodigal' child of Saturn. But converging lines of evidence now suggest that Phoebe formed even farther from the Sun and only later ventured inwards. Its progress was then arrested by gravitational interactions with the Saturn system.

The compositional diversity of Phoebe, in fact, seems unlike that of any single object studied to date within the Solar System. Clark *et al.*⁴ have identified the presence of iron-bearing minerals, and possible phyllosilicates (the family of sheet silicates that includes clays and micas). The latter, if present, would indicate some degree of aqueous processing, possibly in a parent body or even before the formation of the protoplanetary nebula — the cloud of gas and dust from which planets are formed around a newborn star, such as the Sun. Spectral evidence for two crystalline silicate families, the olivines and pyroxenes, is also consistent with the presence of primitive materials left over from the formation of the Solar System.

The spectral observations also point to a wealth of volatile compounds — among them water-ice, carbon dioxide and several organic compounds, including alkanes, aromatic compounds, nitriles and other cyanide compounds — indicating an origin somewhere in the frozen outer reaches of the solar nebula, rather than in the hotter, drier inner Solar System where the terrestrial planets and the asteroid belt formed. An origin beyond Saturn's orbit also makes more sense when considering the orbital dynamics required for outward migration from the asteroid belt, against the gravitational pull of the Sun.

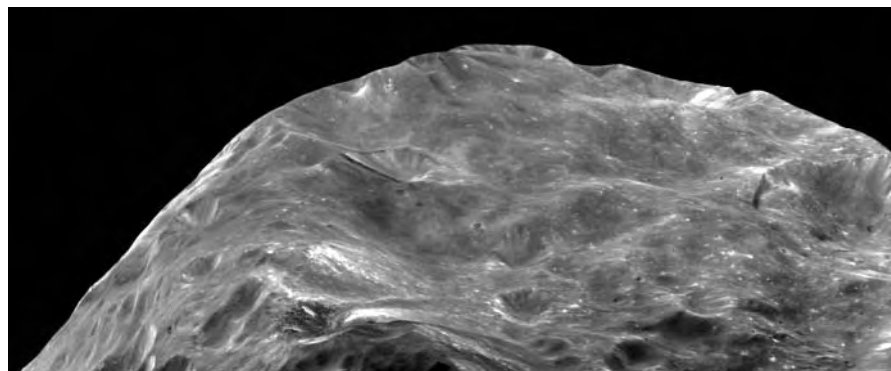


Figure 1 One battered rock. This stunning image of Phoebe's pitted surface, as seen by the Cassini Imaging Science Subsystem, reveals its long and complex history. The region shown is near the satellite's south pole and is around 120 km across; the largest craters are up to 4 km deep. Bright patches of water and other ices, as well as dark regions hosting complex organic compounds and minerals, indicate an origin far from its current orbit around Saturn. Results from Cassini^{3,4} suggest that Phoebe may have formed as far out as the Kuiper belt, beyond Neptune's orbit, and may even contain cometary material pre-dating the formation of the Solar System.

NASA/JPL/SPACE SCIENCE INST.

Box 1 The Cassini–Huygens mission to Saturn

The Cassini–Huygens spacecraft is a joint endeavour of NASA, the European Space Agency and the Italian Space Agency, ASI. With a dry weight of 5.6 tonnes, it is the largest interplanetary spacecraft ever built, and consists of two parts — the orbiter Cassini and the Titan probe Huygens. The components are named after Jean-Dominique Cassini (1625–1711), discoverer of the saturnian moons Dione, Iapetus, Rhea and Tethys, and Christiaan Huygens (1629–95), who first observed Saturn's rings and Titan, its largest moon.

● 15 October 1997: Launch from Cape Canaveral. Helped on its way by 'stealing' rotational energy in three fly-pasts of Venus (April 1998, June 1999) and Earth (August 1999).

● 30 December 2000: Sends back high-resolution images as it passes Jupiter at a distance of 9.7 million kilometres.

● 11 June 2004: First detailed images of Phoebe, on the edge of the Saturn system, from a range of 2,000 kilometres.

● 1 July 2004: Permanently enters orbit around Saturn, sending back spectacular images of the planet's rings.

● 26 October 2004: First close pass of Titan, which, with a diameter of 2,700 km, is the largest of Saturn's 34 known moons.

● 25 December 2004: The Cassini orbiter releases the Huygens probe to coast down to Titan's surface.

● 14 January 2005: Huygens enters Titan's cloudy atmosphere, touching down two-and-a-half hours later. The probe continues sending data for well over an hour from the surface, allowing a detailed picture of its composition to be made. (More will appear on this topic in a later issue.)

In the next three years, Cassini will make numerous further passes of Titan, as well as other moons — including Iapetus, Enceladus, Tethys and Rhea — in the course of 75 orbits of Saturn before its fuel supply is exhausted. **J.B.D.**

The findings of Clark *et al.*⁴ dovetail nicely with Johnson and Lunine's analysis³, which finds Phoebe's mass, volume and density to be more consistent with those of Pluto and Neptune's giant satellite Triton than those of the regular saturnian satellites (with the exception of Hyperion and Titan, which were excluded from the analysis). Because the composition of Phoebe should reflect that of the region in which it formed, this new knowledge could place constraints on the composition and evolution of the early solar nebula, as well as the various worlds spawned within it.

There is much more for Cassini to do. The nature of the proto-saturnian nebula is still not well known; complementary analyses of encounter data for the regular satellites are likely to shed further light on this problem.

The formation of Saturn itself, for example, could have altered the abundances of important species — such as oxygen or water — relative to the protoplanetary nebula⁵. The ring system, as noted by Clark *et al.*, seems to contain similar iron-bearing materials to those seen on Phoebe. It has long been thought that dust from Phoebe could be migrating inwards within the Saturn system, perhaps even providing the bulk of the dark material seen on Saturn's moon Iapetus^{6,7}. The similarities between these two satellites, at least as seen from ground-based telescopes and from the Voyager spacecraft, might also be explained by assuming that meteoritic or cometary infall has coated both with similar material.

The many planned encounters with Iapetus and Saturn's other regular satellites

are expected to provide definitive answers to these remaining questions. Given the far-reaching implications of this first encounter with tiny Phoebe, we can expect Cassini to provide many more powerful insights during its four-year mission. ■

J. Brad Dalton is at the SETI Institute, Mail Stop 245-3, Planetary Systems Branch, NASA Ames Research Center, Moffett Field, California 94035-1000, USA.

e-mail: dalton@mail.arc.nasa.gov

1. Pollack, J. B., Burns, J. A. & Tauber, M. E. *Icarus* 37, 587–611 (1979).
2. Burns, J. A. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 117–158 (Univ. Arizona Press, Tucson, 1986).
3. Johnson, T. V. & Lunine, J. I. *Nature* 435, 69–71 (2005).
4. Clark, R. N. *et al.* *Nature* 435, 66–69 (2005).
5. Mosqueira, I. & Estrada, P. R. *Icarus* 163, 232–255 (2003).
6. Cruikshank, D. P. *et al.* *Icarus* 53, 90–104 (1983).
7. Strazzulla, G. *Icarus* 66, 397–400 (1986).

Reproductive biology

Fatty link to fertility

S. K. Dey

A short delay in the attachment of embryos to the wall of the womb during early pregnancy adversely affects later developmental processes. New evidence reinforces the need for lipids to regulate this event.

In mammals, the creation of new life depends on the union of a sperm and an egg. The fertilized egg then undergoes several cell divisions to form a differentiated tissue — the blastocyst. Next, an intricate molecular dialogue between blastocyst and uterus initiates the process of implantation, by which the blastocyst attaches to the lining of the uterus^{1,2}. Low implantation rates are common in women undergoing assisted reproduction, posing a challenge to both patients and doctors, so researchers are striving towards a deeper understanding of this process. On page 104 of this issue, Ye *et al.*³ reveal a previously unknown and essential role for a small lipid signalling molecule, lysophosphatidic acid (LPA), in normal implantation.

Normal implantation is initiated only when embryonic development is synchronized with the appropriate preparation of the uterus. More specifically, the embryo must reach the blastocyst stage and gain 'implantation competency'; meanwhile, the uterus — through the coordinated actions of oestrogen and progesterone — must reach what is known as the receptive phase.

This state of uterine receptivity, also termed the window of implantation, lasts for a limited period. It is only during this time that the womb is able to support normal embryonic growth and implantation¹. Previous studies have shown that when blastocysts implant beyond this window, the delay creates an adverse ripple effect: embryos crowd near the cervix; placentas fail to form correctly; fetuses are resorbed; and the devel-

opment of remaining fetuses is retarded^{4,5}. In other words, the quality of implantation determines the quality of pregnancy and fetal well-being; failure to achieve 'on-time' implantation risks an adverse pregnancy outcome. In humans, implantation beyond the normal window leads to spontaneous pregnancy losses⁶.

Tracing the hierarchical landscape of the molecular signalling pathways that govern embryo–uterus interactions during human pregnancy is not easy, because of experimental difficulties and ethical considerations. Experiments in mice, however, have directly shown that lipid molecules known as prostaglandins, generated by the enzyme cyclooxygenase-2 (COX-2), are essential for implantation^{5,7}. Their role in reproduction is further illustrated by the poor fertility — caused by deferred implantation — seen in female mice lacking cytoplasmic phospholipase A_{2α} (cPLA_{2α}; ref. 4). This enzyme uses the phospholipids that make up cell membranes to generate arachidonic acid, which is in turn used for prostaglandin synthesis by COX-2. These studies establish the importance of lipid signalling through the cPLA_{2α}–COX-2 axis in implantation.

When a cell is activated in response to a stimulus, membrane phospholipids can be used to generate numerous lipid signalling molecules, such as eicosanoids and lysophospholipids. Prostaglandins belong to the eicosanoid class. The lipid featuring in Ye *et al.*'s study, LPA, belongs to the lysophospholipid group, and is characterized by a 3-carbon backbone, to which is attached an

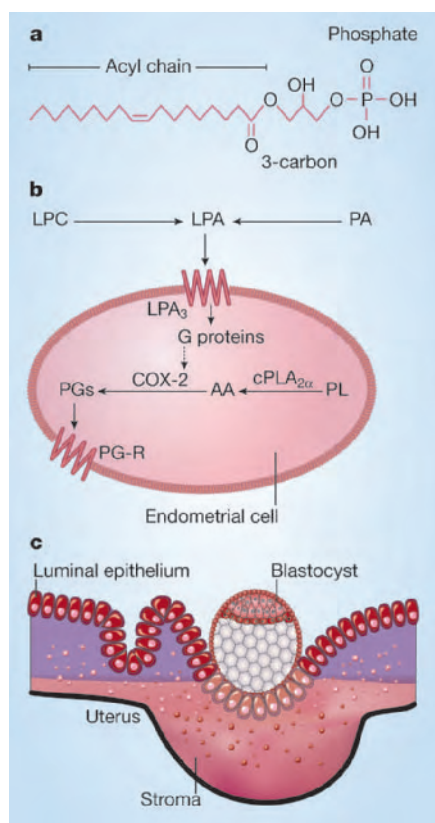


Figure 1 Lipid signals in implantation. Ye *et al.*³ find that lysophosphatidic acid (LPA), through its receptor LPA_3 , is crucial for on-time implantation. **a**, LPA consists of a 3-carbon backbone, attached to an acyl chain and a phosphate head-group. **b**, How LPA and other lipid signals might converge in the uterus. LPA is formed from lysophosphatidylcholine (LPC) or phosphatidic acid (PA). Binding of LPA to LPA_3 in the endometrium (the uterine wall) activates specific G proteins. Ye *et al.* suggest that these in turn activate the enzyme COX-2, leading to greater production of prostaglandins (PGs), which interact with their receptors (PG-Rs). This leads to increased vascular permeability and increased adhesiveness of the uterine lining^{1,2}. Cytosolic $PLA_{2\alpha}$ ($cPLA_{2\alpha}$) is also needed for on-time implantation⁴; it generates arachidonic acid (AA) from membrane phospholipids (PL). **c**, Sites of LPA_3 , COX-2 and $cPLA_{2\alpha}$ expression. LPA_3 is detected in the endometrial luminal epithelium, with peak expression occurring before implantation. $cPLA_{2\alpha}$ and COX-2 are expressed in the endometrial stroma, but also overlap with LPA_3 at the site of blastocyst implantation.

acyl chain and a phosphate head-group (Fig. 1a). Lysophospholipids influence a range of processes — including blood-vessel formation, vascular maturation and neural development — by activating cell-surface receptors that are coupled to common molecular switches termed G proteins⁸. LPA uses four different receptors, LPA_1 to LPA_4 , to produce different effects^{8,9}.

Mice missing LPA_1 and LPA_2 reproduce normally. But Ye *et al.*³ now find that LPA_3 -deficient mice show strikingly similar problems to $cPLA_{2\alpha}$ -deficient ones⁴: deferred implantation, retarded fetal development, embryo crowding, and the sharing of one placenta by several embryos. The result is a much-reduced litter size. Treating these mice with prostaglandins resumes on-time implantation, although embryo crowding persists.

Given the similarities between LPA_3 - and $cPLA_{2\alpha}$ -deficient mice, the reduced levels of uterine COX-2 in mice missing LPA_3 , and the presumed reduction in arachidonic-acid levels (providing less substrate for COX-2) in the $cPLA_{2\alpha}$ -deficient animals, Ye *et al.* conclude that a molecular component that is affected in both cases is COX-2. Thus, both $cPLA_{2\alpha}$ and LPA_3 seem to feed into COX-2 (although how LPA_3 does so is not yet known), thereby influencing the production of prostaglandins and so the timing of implantation (Fig. 1b, c). Ye and colleagues' results also lend further credence to the ripple-effect concept — the idea that embryo implantation occurring past the normal window leads to a poor pregnancy outcome.

It remains to be seen, however, how uterine prostaglandins coordinate with embryonic signals to ensure on-time implantation. Another unresolved issue is the embryo crowding seen in the LPA_3 -deficient mice. This raises questions such as: what signal attracts the embryo to its specific site of attachment in the uterus? Does this finding have any relevance to placenta praevia in humans — an undesirable condition in which the placenta is attached close to or covers the cervix — or to the 'multiple pregnancies' that often arise from assisted reproductive technologies? Although the LPA_3 / $cPLA_{2\alpha}$ -COX-2 signalling axis is

clearly important for proper embryo spacing, we so far have no clue as to how to rescue the crowding defect. The molecules downstream of prostaglandin signalling that participate in the ripple effect also remain unknown.

Defects in blastocyst competency, uterine receptivity and embryo-uterine dialogue all compromise implantation and hence fertility. Mapping the molecular landscape seen during the period of implantation — a period so vital for later development — necessitates well-thought-out experiments to elucidate embryonic and uterine contributions. Past studies have shown a requirement for lipid signalling; Ye and colleagues' work³ now adds LPA signalling to the mix, as an essential player on the maternal side. Ye *et al.* also speculate that still other lipid-based mechanisms may influence fertility, given the reduced fertility, of unknown cause, that is observed when other lysophospholipid receptors are defective¹⁰. Although it is not known whether all these pathways function independently or converge with others to ensure on-time implantation, one thing is clear: fat matters to fertility. ■

S. K. Dey is in the Departments of Pediatrics, Cell & Developmental Biology, and Pharmacology, Division of Reproductive and Developmental Biology, Vanderbilt University Medical Center, 1161 21st Avenue South, Nashville, Tennessee 37232-2678, USA.

e-mail: sk.dey@vanderbilt.edu

1. Dey, S. K. *et al.* *Endocrine Rev.* **25**, 341–373 (2004).
2. Paria, B. C. *et al.* *Science* **296**, 2185–2188 (2002).
3. Ye, X. *et al.* *Nature* **435**, 104–108 (2005).
4. Song, H. *et al.* *Development* **129**, 2879–2889 (2002).
5. Wang, H. *et al.* *J. Biol. Chem.* **279**, 10649–10658 (2004).
6. Wilcox, A. J., Baird, D. D. & Weinberg, C. R. *N. Engl. J. Med.* **340**, 1796–1799 (1999).
7. Lim, H. *et al.* *Cell* **91**, 197–208 (1997).
8. Hla, T. *et al.* *Science* **294**, 1875–1878 (2001).
9. Ishii, I., Fukushima, N., Ye, X. & Chun, J. *Annu. Rev. Biochem.* **73**, 321–354 (2004).
10. Ishii, I. *et al.* *J. Biol. Chem.* **277**, 25152–25159 (2002).

Materials science

A hard look at glass

Paul Madden

The topology of amorphous glasses has generally been considered only at the level of atoms and their nearest neighbours. A larger-scale view gives a fresh perspective on the structure and formation of these glasses.

Glass — window glass being the most familiar example — constitutes an 'ill-condensed' state of matter. Writing on page 75 of this issue, Salmon and colleagues¹ present unprecedentedly detailed studies of the structure of two glassy materials. Their findings may help to explain why the molecules in the materials that form glasses fail to find the lowest-energy arrangement when cooled to low temperatures, and this should aid our understanding of these

scientifically and technologically important materials.

Most liquids, as they are cooled, arrange themselves so as to minimize the interaction energy between their constituent molecules — they crystallize, forming a solid in which the molecules are held in a periodic lattice with long-range order. Distorting this lattice involves sliding whole planes of molecules past each other — giving rise to the familiar rigidity of a perfect crystalline solid.

A glassy material is rigid too, but the arrangement of its molecules is aperiodic ('amorphous'), and does not minimize the energy of the structure². But we can still understand rigidity in amorphous arrangements in the absence of molecular motion (that is, at a temperature of absolute zero) by appealing to mechanical analogues. A random packing of particles with a density above a certain value — a pile of sand, for example — can be rigid. A less-dense collection of particles, in which each element is held to its neighbours by springs, also cannot change its shape if the number of constraints exceeds a critical value. But even with such pictures in mind, fundamental questions remain. Why have amorphous arrangements been adopted at low temperatures in glassy materials? Why not take on the configuration with the lowest energy?

The structures of real (as opposed to computer-modelled) glasses are not known beyond the shortest length scales. Yet it is precisely at longer distances that the consequences of 'ill-condensation' during the cooling process are reflected — and where the answers to these questions, and an understanding of the relationships between different materials, lie.

Salmon *et al.*¹ compare detailed measurements of the atomic arrangements of two simple binary glass-formers — zinc chloride (ZnCl₂) and germanium selenide (GeSe₂) — out to large inter-atomic separations, and demonstrate strong similarities between them. This is remarkable for two reasons. First, the nature of the interatomic interactions in the two materials is expected to differ, because GeSe₂ is more strongly covalent than ZnCl₂. Second, the two materials seem, from the temperature dependence of their viscosities (see ref. 3 and references therein), to approach the transition to the rigid glassy state in characteristically different ways. ZnCl₂ is the less 'fragile' liquid — its approach to rigidity is more continuous than that of GeSe₂, and similar to that of the silica of window glass. One might expect this difference to be correlated with a structural difference in the glasses at longer length scales.

Compared with a crystalline structure, the molecules in a normal liquid will be found at any given instant in a disordered, high-energy configuration — although at the level of the nearest neighbours the structure is almost always similar to that adopted in the solid. At high temperatures, the crystalline state, which minimizes energy, is not favoured because the number of disordered configurations, which give the liquid a higher entropy, is far greater. If a liquid is cooled slowly, the transition between the two states occurs at a sharply defined temperature (the freezing point), where the tendencies to maximize entropy and minimize energy are equal. At this point, groups of molecules adopt the same low-energy structure as the

crystal; once these 'nuclei' exceed a critical size, they can grow spontaneously. The time taken to form such nuclei depends on the rate at which the molecules can reorganize themselves. This rate itself decreases as the temperature of the liquid is lowered — if this is done sufficiently rapidly, the system may reach such low temperatures that reorganization stops before the critical nuclei have a chance to form. The system is left in a rigid, amorphous state — it forms a glass.

This scenario could theoretically occur in any liquid. For a glass to form at practical rates of cooling (those that can be achieved in a laboratory), it is necessary for the formation of the critical nucleus to be 'frustrated', so that the process takes longer than usual. This can be brought about because the preferred local arrangement of the molecules does not readily propagate periodically out to the length scale required to form a crystal nucleus of the critical size. Therefore, experimental attention focuses on the intermediate-range structure of the glass-forming materials, where the frustration of nucleus formation should be apparent.

Salmon and colleagues' experiments are based on neutron diffraction, in which oscillations in the number of neutrons scattering at different angles reflect the relative positions of pairs of atoms. In binary systems such as ZnCl₂ and GeSe₂, the intensity pattern includes information on the distribution of all atom pairs (in the case of zinc chloride, for example, Zn–Zn, Zn–Cl and Cl–Cl). But Salmon *et al.* exploit the fact that neutrons scatter from different isotopes of

the same element with different amplitudes: by performing experiments on three samples with the same chemical but different isotopic compositions, they were able to separate the diffraction pattern connected to each atomic pair and therefore examine the structure of the glass in exquisite detail.

Two characteristic length scales can be distinguished in the observed structures: first, a strong preference for the Zn or Ge atoms to have four Cl or Se nearest neighbours arranged at the local level as a tetrahedron, generating a chemical ordering that extends to large distances; second, an intermediate-range order, associated with the way these tetrahedra are linked together. Despite the differences in the chemistry of ZnCl₂ and GeSe₂, their structures are remarkably similar on both the short and intermediate length scales. Thus, different types of chemical interaction seem to have given rise to similar structures that, from the observation that both materials are good glass-formers, allow for the frustration of crystallization. On the other hand, even at this resolution there is no readily discernible structural feature that can account for the difference in the 'strength' of the two liquids, as seen in the temperature dependence of their viscosity.

Paul Madden is in the School of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JJ, UK.
e-mail: paul.madden@ed.ac.uk

1. Salmon, P. S., Martin, R. A., Mason, P. E. & Cuellar, G. J. *Nature* **435**, 75–78 (2005).
2. Alexander, S. *Phys. Rep.* **296**, 65–236 (1998).
3. Martinez, L. M. & Angell, C. A. *Nature* **410**, 663–667 (2001).

Cell biology

Sterol sensor comes up for air

Renee M. Garza and Randolph Y. Hampton

In one example of a feedback mechanism in mammals, cells switch cholesterol synthesis on or off depending on the availability of sterol. A rewired version of this pathway in yeast acts instead as an oxygen sensor.

Writing in *Cell*, Espenshade and colleagues¹ describe a previously unknown strategy by which cells sense oxygen levels. The mechanism uses an evolutionarily conserved and medically relevant pathway for sterol regulation in an unexpected way.

Over the past few years, a molecular drama has been unfolding in our understanding of how cholesterol synthesis is regulated in mammals. The overarching idea is simple: when cells need more cholesterol, they increase the levels of enzymes that make it, by increasing the expression of the enzyme-encoding genes. But the underlying mechanism for this is quite unexpected. Thanks to heroic work led by Brown and Goldstein², we now have a clear idea of

how cholesterol-regulated gene transcription occurs. Perhaps the only unsurprising feature of this transcriptional regulation is that it centres on a transcription-factor protein: SREBP (for 'sterol-regulatory-element-binding protein').

The SREBP molecule contains a portion that carries out transcription ('TF' in Fig. 1a, overleaf), connected to a transmembrane domain. Freshly made, full-length SREBP is anchored in the membrane of a cellular compartment called the endoplasmic reticulum (ER) — tethered like a pit bull on a chain, and unable to reach its targets in the nucleus. Active SREBP is liberated from its membrane anchor by cleavage, and it is this cleavage that is controlled by sterol levels. Although SREBP resides in the

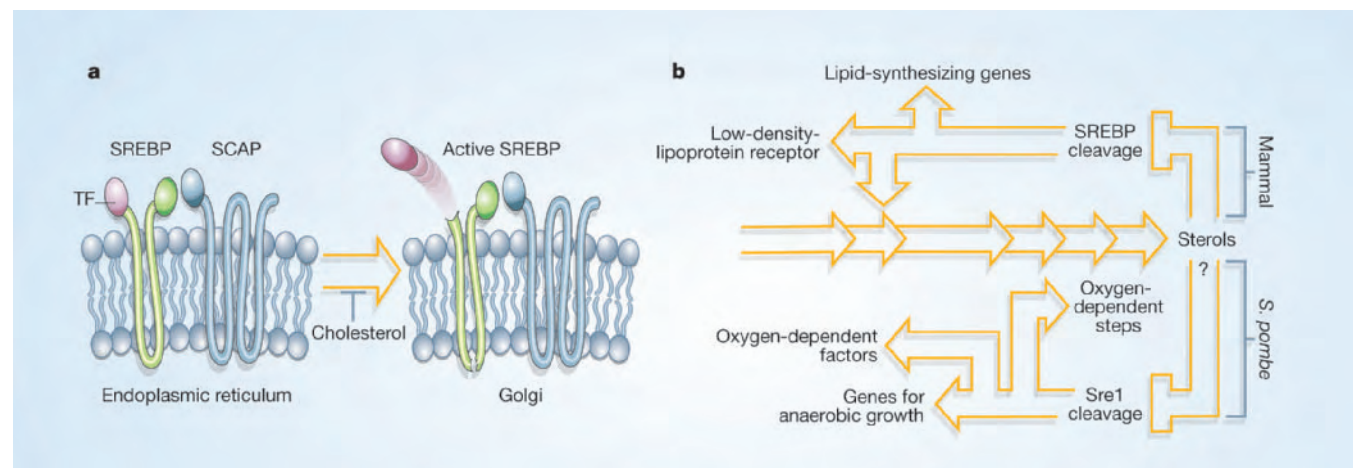


Figure 1 New tricks for a familiar pathway. **a**, The SREBP-processing pathway. The gene-transcription factor SREBP, which regulates sterol biosynthesis, resides in the endoplasmic reticulum in mammalian cells. It is transported to the Golgi by SCAP, where it is cleaved twice to release an active transcription factor (TF). Binding of cholesterol to SCAP inhibits such transport². **b**, Top, mammals use SCAP-dependent cleavage of SREBP to provide feedback control of cholesterol synthesis, and to regulate other

genes involved in lipid metabolism. Thus, if sterol levels are high, SREBP cleavage is inhibited, and these genes are no longer activated. Bottom, the newly discovered 'rewired' pathway from fission yeast¹. The membrane-bound SREBP counterpart Sre1 is regulated by sterols (or possibly other signals) through Scp1, to control genes involved in oxygen-dependent functions and those required for anaerobic growth. This pathway seems to function to sense and respond to oxygen.

ER membrane, the two protein-cleaving enzymes that liberate the active fragment occur in another compartment, the Golgi. When sterol levels are sufficiently low, SREBP is carried from the ER to the Golgi, where cleavage — and thus activation — occurs. When sterol levels are high, the movement of SREBP to the Golgi (and hence cleavage) is blocked.

This 'traffic control' is effected by the protein SCAP (for 'SREBP-cleavage-activating protein'). SCAP binds to SREBP and carries it out of the ER by entering the vesicular pathway that shuttles between the ER and the Golgi. SCAP also has a membrane-embedded motif that directly binds cholesterol. Cholesterol-bound SCAP no longer exits the ER, and so cholesterol inhibits production of active SREBP (Fig. 1a), thereby governing its own synthesis.

Genome sequencing studies show that *Schizosaccharomyces pombe* (fission yeast) also has versions of SREBP and SCAP, called Sre1 and Scp1, respectively. Given that this fungal species synthesizes cholesterol's cousin ergosterol, this is perhaps to be expected. But the surprises came when Espenshade and colleagues¹ explored the molecular details of the *S. pombe* feedback pathway.

The authors found that Sre1 is cleaved to a soluble form and that this cleavage is regulated by sterol in an Scp1-dependent manner, just as in mammals. However, the transcriptional targets of Sre1 are quite distinct. Testing of candidate genes, combined with microarray techniques to analyse gene expression more broadly, showed that the early, rate-limiting reactions of sterol synthesis — those regulated by mammalian SREBP — were completely unaffected by changing levels of active Sre1. Consistent

with this, loss of either Scp1 or Sre1 through mutation had no effect on the growth of *S. pombe*, whereas in mammals loss of the SREBP pathway renders cells totally dependent on added sterols.

Instead, Espenshade and colleagues' analysis indicated that the relevant regulatory function pertains to oxygen. Genes regulated by Sre1 include those encoding: enzymes in the late, oxygen-dependent parts of the sterol-synthesis pathway; enzymes required for the biosynthesis of haem (an oxygen-binding molecule); and several other oxygen-related factors, including some that are needed for yeast cells to survive low oxygen levels. Thus, it seems that sterol-regulated cleavage of Sre1 is used to signal oxygen availability (Fig. 1b).

The authors go on to show that several predictions of their oxygen-sensing model are borne out. First, *S. pombe* cells that lack Sre1 or Scp1 cannot survive in anaerobic conditions. These mutant cells do survive, however, if allowed to express the cleaved form of Sre1 (the 'TF') by molecular-biological trickery. Furthermore, the presence of Sre1 and Scp1 enables the cells to adapt to low-oxygen conditions in ways that would be predicted if these proteins were involved in oxygen sensing — for instance by increasing the cells' capacity for the oxygen-dependent reactions of ergosterol synthesis. Finally, the cleavage of Sre1 is massively stimulated by lowering the oxygen concentration, indicating that the relevant perturbation indeed causes the proposed physiological response.

Although we don't know what signal regulates Scp1-mediated cleavage of Sre1, it is reasonable to suppose that it is a sterol — perhaps ergosterol (the end result of the sterol pathway in *S. pombe*), but possibly some other sterol generated in the pathway,

or a derivative of one of the pathway molecules. And perhaps a second oxygen-dependent signal works together with sterols to regulate Sre1 cleavage. There seems to be great flexibility in the kinds of molecules that can be sensed by the sterol-sensing motif found in SCAP and related proteins^{3–5}, so it may be best to await the next chapter in this new use of the SREBP pathway.

Whatever the specific signals, this 'rewiring' of the SREBP pathway makes a lovely kind of sense. Synthesis of sterols is absolutely dependent on oxygen. Thus, the choice of sterol synthesis as a fiduciary indicator of oxygen levels is a good one, and one that probably exists in more than one yeast species.

This ground-breaking study raises some fascinating questions. Is this a broadly used mode of oxygen sensing in nature? Does it occur in many fungi, and thus provide an Achilles' heel that can be exploited to develop new antifungal drugs? What is the range of molecules sensed by the many SCAP counterparts found in nature? The emerging picture is that these membrane proteins may be widely used sensors of intramembrane signals that affect any aspect of biology in which lipid molecules are involved. ■

Renee M. Garza and Randolph Y. Hampton are in the Section of Cell and Developmental Biology, Division of Biological Sciences, University of California at San Diego, La Jolla, California 92093-0347, USA.

e-mail: rhampton@ucsd.edu

1. Hughes, A. L., Todd, B. L. & Espenshade, P. J. *Cell* **120**, 831–842 (2005).
2. Horton, J. D., Goldstein, J. L. & Brown, M. S. *Cold Spring Harb. Symp. Quant. Biol.* **67**, 491–498 (2002).
3. Shearer, A. G. & Hampton, R. Y. *EMBO J.* **24**, 149–159 (2005).
4. Song, B.-L., Javitt, N. B. & DeBose-Boyd, R. A. *Cell Metabolism* **1**, 179–189 (2005).
5. Dobrosotskaya, I. Y., Seegmiller, A. C., Brown, M. S., Goldstein, J. L. & Rawson, R. B. *Science* **296**, 879–883 (2002).

Plant biology

Chestnuts clock out in winter

Proc. Natl Acad. Sci. USA
doi:10.1073/pnas.0408549102 (2005)

Do plants put their internal clocks on pause when winter arrives? According to Alberto Ramos *et al.*, such a disruption indeed takes place in the European chestnut.

Ramos *et al.* examined parts of adult trees — including stems, leaves and buds — for gene products associated with the circadian clock. They found that the chestnut clock behaves like that of the flowering plant *Arabidopsis thaliana* during the growing season, but acts differently in colder months.

During June, expression of the genes *CsTOC1* and *CsLHY* (counterparts of two *Arabidopsis* clock genes) followed strong cycles. But their expression remained high and did not oscillate when the trees encountered colder temperatures and entered dormancy. Normal cycling of *CsTOC1* and *CsLHY* expression resumed when chestnut seedlings in the lab were exposed to warmer temperatures. The findings suggest that woody and herbaceous plants use different methods for adjusting their clocks to the cold.

Roxanne Khamsi

Cell biology

Prion progress

Cell **121**, 195–206 (2005)

Are prions — infectious proteins — the sole cause of disorders such as Creutzfeldt–Jakob disease and scrapie? Joaquín Castilla *et al.* add to the case that they are.

According to the prion hypothesis, normal prion proteins become twisted out of shape, and then convert further normal proteins into the same warped form. One way of testing the hypothesis is to generate misshapen prions *in vitro* and see if they cause disease *in vivo*. Researchers have succeeded in doing this with yeast prions. They have also generated mammalian prion-like protein fragments *in vitro*, and used them to cause disease in transgenic mice engineered to overexpress the same fragments.

Castilla *et al.* go further. Tweaking one of their own techniques, called ‘protein misfolding cyclic amplification’, they used extracts of scrapie-infected hamster brains to ‘seed’ the conversion of normal prion proteins into the abnormal form *in vitro*. Numerous cycles produced samples containing practically none of the original brain extracts. When wild-type hamsters were infected with prions that had been generated *in vitro*, the animals showed typical signs of scrapie.

Amanda Tromans

Network theory

On Broadway

Science **308**, 697–702 (2005)

Another opening, another show. But what makes it a success? According to Roger Guimerà and colleagues’ analysis of Broadway musicals, the answer is maximizing experience and innovation by means of a well-mixed and well-connected production team, with a dash of fresh faces.

Underlying such teams is a network of collaborations, and Guimerà *et al.* investigate these in terms of team size, which grows with the complexity of the task; the probability of being an established team member, or ‘incumbent’, hence experienced and well connected; and the

likelihood that an incumbent chooses to work again with a previous collaborator.

Simulating team-building in the phase space determined by these parameters, the authors find a phase transition to a large connected cluster. This means that there are optimum values of the parameters for which a complex network of contacts exists that is the basis for success.

The same is true of some research teams. For collaborations in social psychology, economics and ecology — and using journal impact factor as a measure of success — Guimerà *et al.* conclude that it’s getting the right combination of newcomers and incumbents that counts. Curiously, however, astronomy doesn’t fit the pattern.

Alison Wright

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Chemistry

Molecular bends

Angew. Chem. Int. Edn Engl. **44**, 2382–2385 (2005)

Heating usually speeds up chemical reactions — the extra energy causes molecules to vibrate more, which stretches interatomic bonds and makes them more likely to break. Hans A. Bechtel and colleagues find that heating also increases the bending of molecular bonds.

Compared with stretching vibrations, bending vibrations have a lower frequency and are less energetic, so have remained largely unexplored. Bechtel *et al.* studied the reaction of methane (CH_4) with atomic chlorine, which occurs in atmospheric chemistry and is a common model for investigating reaction rates.

They show that, counterintuitively, bending excitation increases the chances of methane reacting by at least a factor of two, implying that shearing motion, as well as stretching, can help to break C–H bonds. They also find that the energy involved in the movement of methane’s atoms leads almost exclusively to motion of the escaping reaction products, rather than being retained as vibrational energy in the products — something not predicted by theory. Targeting bending excitation of molecules may boost rates in other reactions, the authors suggest.

Mark Peplow

Drug delivery

Shake-up for injections

Colloids Surf. A **260**, 7–16 (2005)

Many drugs are built on a hydrocarbon molecular scaffolding, and so are hydrophobic and rather insoluble in water. This makes intravenous injection of water-based solutions impossible. Often, such drugs are dissolved in oils that are then dispersed in water using surfactant emulsifiers — but both the oils and the surfactants may degrade into harmful compounds.

R. M. Pashley and co-workers recently found that water and oil mixtures from which dissolved air has been removed are much better at dispersing oily and greasy substances. Pashley and M. J. Francis have now shown that degassed oils dispersed in degassed water could be a suitable method for delivering hydrophobic drugs.

Simply shaking common drug-delivery oils, in degassed form, such as soybean oil, degassed water for a few seconds produces a dispersion of uniform droplets about 1–3 μm in diameter, which could act as robust ‘parcels’ for the safe intravenous injection of drugs. Liquid, water-insoluble drugs such as propofol (a sedative) and griseofulvin (used to treat skin infections) could be dispersed directly in degassed water without any carrier oil.

Philip Ball

Specialized bird perch aids cross-pollination

A plant scores by providing an access point for visiting sunbirds to feed on its nectar.

Birds may hover over or perch on flowers when feeding on nectar¹, and this assists cross-pollination if they then visit other plants. Here we investigate the curious sterile inflorescence axis of the South African Cape endemic 'rat's tail' plant (*Babiana ringens*, Iridaceae), whose function — unlike in other bird-pollinated plants — is exclusively to provide a perch for foraging birds. We find that this structure promotes the plant's mating success by causing the malachite sunbird (*Nectarinia famosa*), its main pollinator, to adopt a position ideal for the cross-pollination of its unusual ground-level flowers.

The Cape naturalist Rudolf Marloth was the first to propose that the rat's tail of *B. ringens* (Fig. 1a) could function as a perch to facilitate cross-pollination by visiting sunbirds (cited in ref. 2). We investigated this proposal in two populations of *B. ringens* near Mamre, Western Cape Province, from August to October in 2003 and 2004. Our observations indicated that the plants' only pollinator was the malachite sunbird, the largest of the sunbirds that visit species of Iridaceae in the Cape³.

We observed birds alighting on the sterile axis of the plant and rotating upside down before probing the flowers for nectar (Fig. 1b). Because the floral tube in *B. ringens* curves upwards, rather than downwards as in most bird-pollinated species, a sunbird accesses nectar by inserting its curved beak from above. The plant's sexual organs contact the bird's breast, enabling pollen to be transferred between plants.

To test the function of the bird perch, we compared fitness components in unmanipulated plants with those from which we had removed the sterile axis just before flowering. Perch removal did not alter the floral display or the intrinsic ability of plants to set seed. We found no significant difference in seed set between plants with and without perches following supplemental cross-pollination (repeated measures analysis of variance, specific contrast: $t = -0.87$, $P = 0.39$). Perch removal therefore does not interfere with seed provisioning, and the perch is not a significant source of photosynthetic carbon for seed development. (For methods, see supplementary information.)

Although perch removal did not preclude visitation by sunbirds (Fig. 1c), they preferred plants with intact perches. Of 93 visits recorded after perches were removed from half the plants, 59 were made to plants with perches and 34 were made to plants without perches ($G = 6.80$, $P = 0.009$). Female and male birds showed different preferences for

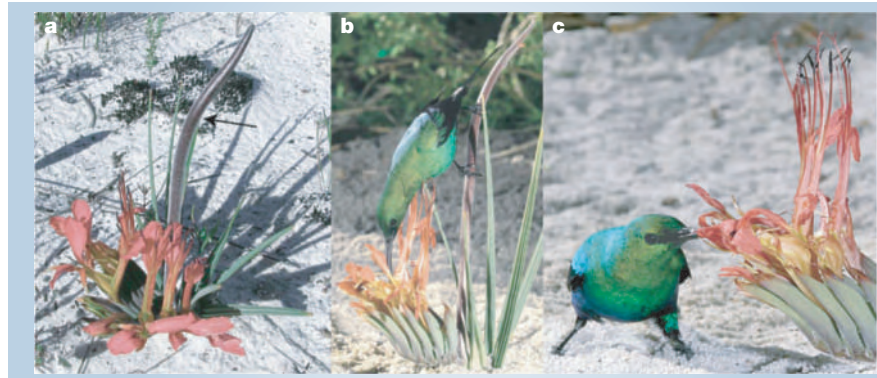


Figure 1 Features of the South African plant *Babiana ringens*. **a**, The specialized bird perch (arrow) and ground-based flowers of *B. ringens*. **b, c**, Male sunbirds (*Nectarinia famosa*) visiting the plant adopt different positions to feed according to whether the perch is **b**, present, or **c**, absent.

plants with and without perches: only males had a strong preference for plants with perches. Males visited significantly more plants (total visits by males: perch present, 24; perch removed, 11; $G = 4.95$, $P = 0.026$; total visits by females: perch present, 35; perch removed, 23; $G = 2.50$, $P = 0.113$; NS). Males spent longer feeding from flowers (time spent foraging by males: perch present, 53.95 s (s.e. 9.42); perch removed, 12.4 s (s.e. 4.72) (paired t -test, $t = 3.92$, d.f. = 19, $P = 0.0005$); time spent foraging by females: perch present, 23.88 s (s.e. 4.87); perch removed, 13.34 s (s.e. 4.18) ($t = 1.52$, d.f. = 31, $P = 0.07$)).

This sex difference in preference may be associated with the much longer tail feathers of males, which could interfere with ground landings and suffer damage. Males may also be more reluctant to land because of the predation risk associated with their bright colours, or because perching in prominent places may be linked to territoriality.

Perch removal reduced female fertility: plants without perches produced 47% fewer seeds on average than unmanipulated plants (Fig. 2a). Plants without perches set considerably more seed than bagged inflorescences, however, so sunbird-mediated self-pollination seems likely to have occurred in these plants. We confirmed this by analysing mating patterns. Perch removal increased self-fertilization by an average of 35% (Fig. 2b), indicating that perch specialization reduces the genetic costs of self-pollination⁴.

Fitness advantages to mating and fertility arise because the perch manipulates the position of foraging sunbirds for better pollen dispersal. The specialized bird perch of *B. ringens* is a novel example of a structural adaptation that promotes cross-pollination in angiosperms.

Bruce Anderson*, **William W. Cole†**,
Spencer C. H. Barrett†

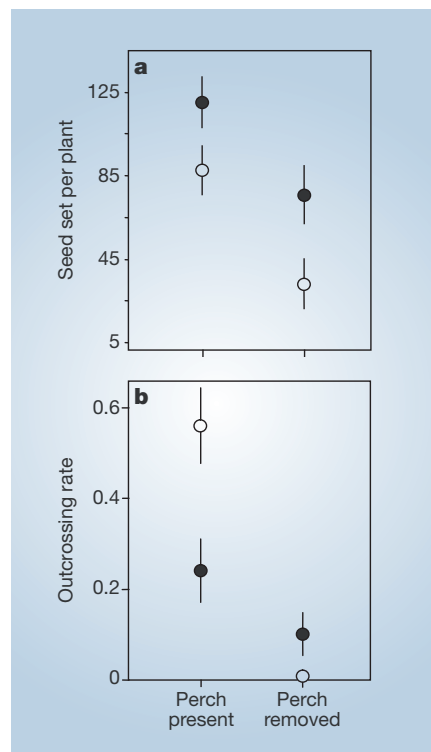


Figure 2 The specialized bird perch of *Babiana ringens* increases mating success. A fraction of each of two plant populations (A and B, represented by filled and open circles) located about 5 km apart had the bird perch removed (in 11 of 25 and 19 of 38 plants, respectively) and the seed set and outcrossing rate were compared with those of unmanipulated plants. **a**, Perch removal decreases seed production per plant (treatment: $F_{1,52} = 15.05$, $P < 0.001$; no significant treatment $\times$ population interaction: $F_{1,52} = 0.16$, $P = 0.692$). **b**, Perch removal increases the mean ($\pm 95\%$ confidence interval) frequency of self-fertilization (population A, $P = 0.019$; population B, $P < 0.001$). Genotype scoring at five allozyme loci was used to determine the outcrossing rates, with the following sample sizes for each population: families from A, 25; families from B, 30; seeds per family from A, 8.4; seeds per family from B, 8.5. A family is the collection of seeds from one maternal parent. For details, see supplementary information.

*School of Botany and Zoology, University of KwaZulu-Natal, Private Bag X01, Scottsville, Pietermaritzburg 3209, South Africa

†Department of Botany, University of Toronto, Toronto, Ontario M5S 3B2, Canada
e-mail: barrett@botany.utoronto.ca

1. Westerkamp, C. *Bot. Acta* **103**, 366–371 (1990).
2. Brown, N. E. *Trans. R. Soc. S. Africa* **20**, 261–280 (1932).
3. Goldblatt, P., Manning, J. C. & Bernhardt, P. *Adansonia* **21**, 25–40 (1999).
4. Barrett, S. C. H. & Harder, L. D. *Trends Ecol. Evol.* **11**, 73–79 (1996).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Biochemistry

A cadmium enzyme from a marine diatom

The ocean biota contains a vast reservoir of genomic diversity¹. Here we present the sequence and preliminary characterization of a protein that is a cadmium-containing carbonic anhydrase from the marine diatom *Thalassiosira weissflogii*. The existence of a cadmium enzyme in marine phytoplankton may indicate that there is a unique selection pressure for metalloenzymes in the marine environment², and our discovery provides a long-awaited explanation for the nutrient-like behaviour of cadmium in the oceans³.

In marine diatoms, the metals cadmium, cobalt and zinc can functionally substitute for one another to maintain optimal growth rates. This effect is at least partly due to metal replacement in the metal-binding site of the enzyme carbonic anhydrase⁴, which is involved in the acquisition of inorganic carbon for photosynthesis. In addition to a zinc carbonic anhydrase that can substitute cobalt in its active site *in vivo*⁵, *T. weissflogii* has a putative cadmium carbonic anhydrase that is also involved in acquiring inorganic carbon⁶.

We purified a protein, CDCA1, from this organism that has carbonic anhydrase activity and contains cadmium (Fig. 1a, b; Genbank accession number AY772014). Determination of its sequence was complicated by the presence of a triple repeat (see supplementary information): the three amino-acid sequences are virtually identical (about 85% identity), but there is more variation in their encoding DNA (about 78% identity). Contrary to an earlier estimate⁶, we have determined the relative molecular mass of CDCA1 as about 69K (for methods, see supplementary information). CDCA1 has probably not been sequenced before, as there are no hits in the NCBI database. It is significantly different from any of the known major classes of carbonic anhydrases^{7,8}, and therefore represents the first member of a new class of carbonic anhydrases, the ζ class.

The genome of *T. pseudonana* contains a single sequence that is highly homologous to

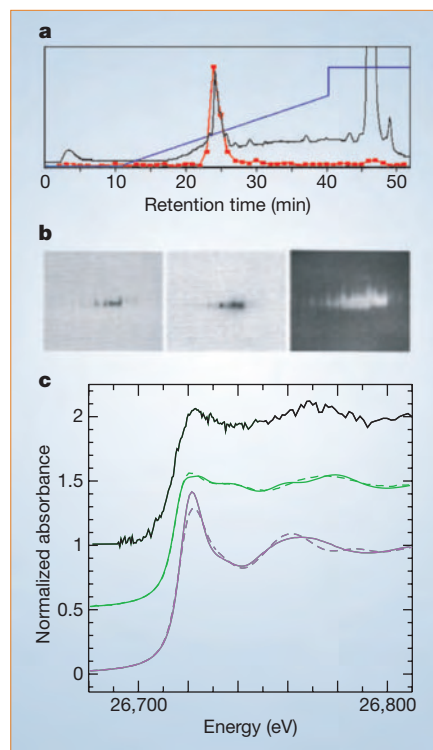


Figure 1 Characterization of the carbonic anhydrase CDCA1 from a marine diatom. **a**, High-pressure liquid chromatography (HPLC) of the CDCA1 enzyme, showing co-elution of purified CDCA1 (black trace, absorbance at 280 nm) and of ¹⁰⁹Cd-radiolabelled enzyme (red, c.p.m., 330 c.p.m. at peak) at 24 min. Blue line, NaCl gradient. **b**, Non-denaturing gel electrophoresis of HPLC fractions containing the main peak of protein and radiolabel shows an intense Coomassie-blue-stained protein band (left) that co-migrates with the ¹⁰⁹Cd radiolabel (centre) and carbonic anhydrase activity (right). **c**, Cadmium K-edge X-ray absorption spectra from purified CDCA1 (top) compared with two tetrahedral, thiolate-coordinated species, [Cd(SPh)₄](Me₆N)₂ (green solid line), cadmium phytochelatin (green broken line), and with two octahedral species [Cd(H₂O)₆]²⁺ (purple solid line) and [Cd(midazole)₆](NO₃)₂ (purple broken line). The similarity of the CDCA1 spectrum to that of the tetrahedral species suggests that the metal site has tetrahedral symmetry and involves cysteinyl ligands to the metal. (For methods, see supplementary information.)

the three repeats in *T. weissflogii*, corresponding to a protein of 25.5K (see supplementary information). The presence of expressed-tag sequences shows that this putative cadmium-containing carbonic anhydrase is expressed, indicating that only a single repeat of CDCA1 may be necessary for activity. The amino-terminal sequence of the *T. pseudonana* enzyme (not available for CDCA1) has only 15 amino acids preceding the homologous sequence from *T. weissflogii*.

X-ray absorption near-edge spectroscopy of the purified protein confirms the presence of the cadmium-binding site (Fig. 1c). Comparison with standards⁹ indicates that the site probably has a roughly tetrahedral geometry and that the cadmium ion is bound by two or more thiolates — as in the zinc-containing β class of carbonic anhydrases in higher plants, which have two cysteines and a histidine at the metal-binding site¹⁰. The spectra are also consistent with an

active site containing an activated water molecule, as in other carbonic anhydrases.

Gene-expression analysis shows that there is an increase in the abundance of *cdca1* transcripts within 1 hour of increasing the concentration of cadmium or of decreasing the partial pressure of carbon dioxide in the medium (results not shown). Expression of CDCA1 may therefore be controlled in part by the availability of cadmium and carbon dioxide in sea water.

High-throughput sequencing of a sea-water sample has revealed that the marine environment may contain unique genes¹. We have identified and partially characterized one such gene — for a carbonic anhydrase from a marine diatom that, to our knowledge, is the first native enzyme so far discovered to contain cadmium. Because of the extraordinarily low concentrations of many essential trace metals in sea water, it is likely that there are other metalloenzymes in marine organisms that use unusual metals for activity and contribute to trace-metal geochemical cycling in the oceans.

Todd W. Lane*, Mak A. Saito†, Graham N. George‡, Ingrid J. Pickering‡, Roger C. Prince§, François M. M. Morell||

*Biosystems Research Department, Sandia National Laboratories, Livermore, California 94551-0969, USA

†Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

e-mail: mak@whoi.edu

‡Department of Geological Sciences, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5E2, Canada

§ExxonMobil Research and Engineering Company, Annandale, New Jersey 08801, USA

||Department of Geosciences, Princeton University, Princeton, New Jersey 08544, USA

1. Venter, J. C. *et al. Science* **304**, 66–74 (2004).
2. Saito, M. A., Sigman, D. & Morel, F. M. M. *Inorg. Chim. Acta* **356**, 308–318 (2003).
3. Boyle, E. A., Slater, F. R. & Edmond, J. M. *Nature* **263**, 42–44 (1976).
4. Price, N. M. & Morel, F. M. M. *Nature* **344**, 658–660 (1990).
5. Morel, F. M. M. *et al. Nature* **369**, 740–742 (1994).
6. Lane, T. W. & Morel, F. M. M. *Proc. Natl Acad. Sci. USA* **97**, 4627–4631 (2000).
7. Smith, K. S. & Ferry, J. G. *FEMS Microbiol. Rev.* **24**, 335–366 (2000).
8. So, A. K. C. *et al. J. Bacteriol.* **186**, 623–630 (2004).
9. Pickering, I. J. *et al. Biochim. Biophys. Acta* **1429**, 351–364 (1999).
10. Kimber, M. S. & Pai, E. F. *EMBO J.* **19**, 1407–1418 (2000).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Corrigendum

Animal behaviour: Elephants are capable of vocal learning

J. H. Poole, P. L. Tyack, A. S. Stoeger-Horwath, S. Watwood *Nature* **434**, 455–456 doi:10.1038/434455a (2005)

The African elephant Calimero spent 18 years with two Asian elephants at Bioparco in Rome, Italy, not the Basel Zoo, Switzerland. He was later transferred to Basel, where the recordings described in this Brief Communication were made.

The genome of the social amoeba *Dictyostelium discoideum*

L. Eichinger^{1*}, J. A. Pachebat^{1,2*}, G. Glöckner^{3*}, M.-A. Rajandream^{4*}, R. Sugang^{5*}, M. Berriman⁴, J. Song⁵, R. Olsen⁹, K. Szafranski³, Q. Xu^{6,7}, B. Tunggal¹, S. Kummerfeld², M. Madera², B. A. Konfortov², F. Rivero¹, A. T. Bankier², R. Lehmann³, N. Hamlin⁴, R. Davies⁴, P. Gaudet¹⁰, P. Fey¹⁰, K. Pilcher¹⁰, G. Chen⁵, D. Saunders⁴, E. Sodergren^{6,8}, P. Davis⁴, A. Kerhornou⁴, X. Nie⁵, N. Hall^{4,†}, C. Anjard⁹, L. Hemphill⁵, N. Bason⁴, P. Farbrother¹, B. Desany⁵, E. Just¹⁰, T. Morio¹¹, R. Rost¹², C. Churcher⁴, J. Cooper⁴, S. Haydock¹³, N. van Driessche⁶, A. Cronin⁴, I. Goodhead⁴, D. Muzny⁸, T. Mourier⁴, A. Pain⁴, M. Lu⁵, D. Harper⁴, R. Lindsay⁵, H. Hauser⁴, K. James⁴, M. Quiles⁸, M. Madan Babu², T. Saito¹⁴, C. Buchrieser¹⁵, A. Wardroper^{2,16}, M. Felder³, M. Thangavelu¹⁷, D. Johnson⁴, A. Knights⁴, H. Louiseged⁸, K. Mungall⁴, K. Oliver⁴, C. Price⁴, M. A. Quail⁴, H. Urushihara¹¹, J. Hernandez⁸, E. Rabinowitsch⁴, D. Steffen⁸, M. Sanders⁴, J. Ma⁵, Y. Kohara¹⁸, S. Sharp⁴, M. Simmonds⁴, S. Spiegler⁴, A. Tivey⁴, S. Sugano¹⁹, B. White⁴, D. Walker⁴, J. Woodward⁴, T. Winckler²⁰, Y. Tanaka¹¹, G. Shaulsky^{6,7}, M. Schleicher¹², G. Weinstock^{6,8}, A. Rosenthal³, E. C. Cox²¹, R. L. Chisholm¹⁰, R. Gibbs^{6,8}, W. F. Loomis⁹, M. Platzer³, R. R. Kay², J. Williams²², P. H. Dear², A. A. Noegel¹, B. Barrell⁴ & A. Kuspa^{5,6}

¹Center for Biochemistry and Center for Molecular Medicine Cologne, University of Cologne, Joseph-Stelzmann-Str. 52, 50931 Cologne, Germany

²Laboratory of Molecular Biology, MRC Centre, Cambridge CB2 2QH, UK

³Genome Analysis, Institute for Molecular Biotechnology, Beutenbergstr. 11, D-07745 Jena, Germany

⁴The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK

⁵Verna and Marrs McLean Department of Biochemistry and Molecular Biology, ⁶Department of Molecular and Human Genetics, ⁷Graduate Program in Structural and Computational Biology and Molecular Biophysics, and ⁸Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA

⁹Section of Cell and Developmental Biology, Division of Biology, University of California, San Diego, La Jolla, California 92093, USA

¹⁰dictyBase, Center for Genetic Medicine, Northwestern University, 303 E Chicago Ave, Chicago, Illinois 60611, USA

¹¹Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan

¹²Adolf-Butenandt-Institute/Cell Biology, Ludwig-Maximilians-University, 80336 Munich, Germany

¹³Biochemistry Department, University of Cambridge, Cambridge CB2 1QW, UK

¹⁴Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

¹⁵Unité de Genomique des Microorganismes Pathogènes, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cedex 15, France

¹⁶Department of Biology, University of York, York YO10 5YW, UK

¹⁷MRC Cancer Cell Unit, Hutchison/MRC Research Centre, Hills Road, Cambridge CB2 2XZ, UK

¹⁸Centre for Genetic Resource Information, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan

¹⁹Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Minato, Tokyo 108-8639, Japan

²⁰Institut für Pharmazeutische Biologie, Universität Frankfurt (Biozentrum), Frankfurt am Main 60439, Germany

²¹Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544-1003, USA

²²School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK

* These authors contributed equally to this work

† Present address: The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA

The social amoebae are exceptional in their ability to alternate between unicellular and multicellular forms. Here we describe the genome of the best-studied member of this group, *Dictyostelium discoideum*. The gene-dense chromosomes of this organism encode approximately 12,500 predicted proteins, a high proportion of which have long, repetitive amino acid tracts. There are many genes for polyketide synthases and ABC transporters, suggesting an extensive secondary metabolism for producing and exporting small molecules. The genome is rich in complex repeats, one class of which is clustered and may serve as centromeres. Partial copies of the extrachromosomal ribosomal DNA (rDNA) element are found at the ends of each chromosome, suggesting a novel telomere structure and the use of a common mechanism to maintain both the rDNA and chromosomal termini. A proteome-based phylogeny shows that the amoebozoa diverged from the animal–fungal lineage after the plant–animal split, but *Dictyostelium* seems to have retained more of the diversity of the ancestral genome than have plants, animals or fungi.

The amoebozoa are a richly diverse group of organisms whose genomes remain largely unexplored. The soil-dwelling social amoeba *Dictyostelium discoideum* has been actively studied for the past 50 years and has contributed greatly to our understanding of cellular motility, signalling and interaction¹. For example, studies in *Dictyostelium* provided the first descriptions of a eukaryotic cell chemoattractant and a cell–cell adhesion protein^{2,3}.

Dictyostelium amoebae inhabit forest soil and consume bacteria and yeast, which they track by chemotaxis. Starvation, however, prompts the solitary cells to aggregate and develop as a true multicellular organism, producing a fruiting body comprised of a cellular, cellulosic stalk supporting a bolus of spores. Thus, *Dictyostelium* has evolved mechanisms that direct the differentiation of a homogeneous population of cells into distinct cell types, regulate

the proportions between tissues and orchestrate the construction of an effective structure for the dispersal of spores⁴. Many of the genes necessary for these processes in *Dictyostelium* were also inherited by Metazoa and fashioned through evolution for use within many different modes of development.

The amoebozoa are also noteworthy as representing one of the earliest branches from the last common ancestor of all eukaryotes. Each of the surviving branches of the crown group of eukaryotes provides an example of the ways in which the ancestral genome has been sculpted and adapted by lineage-specific gene duplication, divergence and deletion. Comparison between representatives of these branches promises to shed light not only on the nature and content of the ancestral eukaryotic genome, but on the diversity of ways in which its components have been adapted to meet the needs

of complex organisms. The genome of *Dictyostelium*, as the first free-living protozoan to be fully sequenced, should be particularly informative for these analyses.

Mapping, sequencing and assembly

An international initiative to sequence the genome of *Dictyostelium discoideum* AX4 (refs 5, 6) was launched in 1998. The high repeat content and (A+T)-richness of the genome (the latter rendering large-insert bacterial clones unstable) posed severe challenges for sequencing and assembly. The response to these challenges was to use a whole-chromosome shotgun (WCS) strategy, partially purifying each chromosome electrophoretically and treating it as a separate project. This approach was supported by novel statistical tools to recover chromosome specificity from the impure WCS libraries, and by highly detailed HAPPY maps that provided a framework for sequence assembly. These approaches have enabled the completion of this difficult genome to a high standard, and are likely to be valuable in tackling the many other genomes that present challenges of composition and complexity.

Genome mapping

To support sequence assembly, we made high-resolution maps of the chromosomes using HAPPY mapping^{7–9}, which relies on analysing the sequence content of single DNA molecules prepared by limiting dilution. A total of 3,902 markers selected mostly from the emerging shotgun data were mapped, and maps of all six chromosomes were assembled (see Methods and Table 1; see also Supplementary Fig. 1 and Supplementary Table 1).

Genome sequencing and assembly

Two strategies were used to recover chromosome-specific data from impure WCS libraries (see Methods). The first (for chromosomes 1, 2 and 3) used enrichment of the respective libraries as the main statistical indicator of the chromosomal assignment of contigs, and HAPPY maps were used to guide assembly. The second strategy (for chromosomes 4, 5 and most of 6) used mapping data to assign sequences to chromosomes initially, with detailed HAPPY maps being used to validate final assemblies. A 1,508-kilobase (kb) portion of chromosome 6 was sequenced as a pilot project using a combination of approaches (see Methods).

Repetitive tracts complicated assembly. For chromosomes 1, 2 and 3, inspection of polymorphisms, combined with HAPPY maps, allowed unambiguous assembly in many cases. For chromosomes 4,

5 and 6, low-coverage sequencing of AX4-derived yeast artificial chromosomes (YACs) alleviated the problems by providing a local data set within which the troublesome repeat element was present as a single copy. Nevertheless, some repeat tracts proved intractable and remain as gaps. Thirty-four unlinked (floating) contigs of >1 kb, totalling 225,339 base pairs (bp), remain unpositioned in the genome, but can be provisionally assigned to specific chromosomes based on their content of reads from the WCS libraries. Most or all of these floating contigs are bounded by repetitive regions. The chromosome 2 sequence in the current assembly supersedes that previously published⁹, having benefited from further HAPPY mapping and manual sequence finishing.

The six chromosomal assemblies span 33,817 kb (Table 1), including ~156 kb in the form of clone-, sequence- and repeat gaps. Assuming that most of the floating contigs lie beyond the termini of the assemblies, the total genome size is estimated at 34,042,810 bp. In estimating the completeness of the sequence, we note that of 967 well-characterized *D. discoideum* genes, 957 (99%) were found initially in the assemblies. Of the remaining ten, seven (*cupE*, *trxA*, *trxB*, *trxC*, *staA*, *staB* and *cinB*) have close matches, suggesting that their GenBank entries may contain errors or represent alternative alleles. Only three (*fcpA*, *wasA* and *roco5*) had no matches in the initial assemblies, although the first two of these were recovered by searches of unincorporated sequence followed by local reassembly. Of 133,168 'qualified' *D. discoideum* AX4 expressed sequence tags (ESTs of >200 bp and >20% G+C, and not matching mitochondrial sequence; ref. 10 and H. Urushihara *et al.*, unpublished data), 128,207 (96.3%) are found in the assemblies (the higher proportion of missing sequences among the ESTs probably reflects the higher error rate inherent in EST data).

We conclude that the current assembly represents >>95% of the chromosomal sequence (less than 1% of which is in floating contigs) and ≥99% of genes, with most of the missing sequence comprising complex or simple repeats. The most stringent test of the medium-to long-range accuracy of the assembly comes from comparison with the HAPPY maps. This is particularly true for chromosomes 4, 5 and 6, where HAPPY markers were used to nucleate contigs but not to guide their assembly or ordering, specifically to allow such a comparison to be made without circularity of argument. As can be seen, good agreement between map and sequence confirms the accuracy of the assembly (Fig. 1).

Table 1 Sequence assembly details

Feature	Chromosome						All
	1	2	3	4	5	6	
Chromosomal assemblies							
Assembly span (bp)*	4,919,822	8,467,571	6,358,352	5,430,575	5,062,323	3,578,828	33,817,471
Assembly sequence (bp)†	4,911,622	8,437,971	6,334,852	5,397,875	5,032,273	3,547,128	33,661,721
Total contigs	11	40	32	65	107	44	309
Mean contig size (bp)	446,511	210,949	197,964	83,044	47,031	80,617	108,938
Number of sequence gaps	4	12	10	34	81	14	155
Number of repeat gaps	8	29	23	9	4	11	84
Number of clone gaps	0	0	0	22	22	20	64
Total estimated gap size (bp)‡	8,200	29,600	23,500	32,700	30,050	31,700	155,750
Number of HAPPY markers (mean spacing in kb)	749 (6.6)	615 (12.5)§	684 (9.3)	628 (8.6)	628 (8.1)	598 (6.0)	3,902 (8.7)
Floating contigs							
Number of floating contigs	0	22	3	9¶		0	34
Total size of floating contigs (bp)	0	171,670	16,360	37,309		0	225,339
Combined (assemblies plus floating contigs)							
Total sequence (bp)	4,911,622	8,609,641	6,351,212	5,416,529¶	5,050,928¶	3,547,128	33,887,060
Mean coverage (fold)	9.1	6.5	6.7	9.6	9.9	10.3	8.3

*Total end-to-end length of the chromosomal assembly, including any gaps.

†Sequenced bases covered by chromosomal assembly, not counting gaps.

‡Sequence, repeat and clone gaps are taken to have average sizes of 50 bp, 1,000 bp and 1,000 bp, respectively.

§Does not include the second copy of the 755-kb inverted duplication.

||Includes only those contigs that can be assigned to specific chromosomes.

¶Floating contigs from chromosomes 4 and 5 cannot be distinguished. In calculating total chromosomal sequence, we assume that half of these floating contigs are from each of chromosomes 4 and 5.

Sequence characteristics of the genome

The genome is (A+T)-rich (77.57%) and has a broadly uniform composition, apart from the more (G+C)-rich repeat-dense regions (Fig. 2). On a finer scale, nucleotide composition tracks the distribution of exons (see below). Among dinucleotides, CpG is under-represented, not just in absolute terms but also relative to its isomer GpC (the former occurring only 62% as often as the latter). This bias normally reflects cytosine methylation at CpG sequences, promoting their mutation to TpG (which is over-represented relative to GpT by 38%). Hence, these observations suggest that cytosine methylation may occur in *Dictyostelium*, contrary to earlier findings¹¹.

Simple sequence repeats are abundant and unusual

Simple sequence repeats (SSRs) are more abundant in *Dictyostelium* than in any other genome sequenced so far, comprising >11% of bases (Supplementary Fig. 2). In non-coding sequence, tracts of dinucleotides or longer motifs occur every 392 bp on average and comprise 6.4% of the bases. There is a bias towards repeat units of 3–6 bases, whereas dinucleotide tracts predominate in most other genomes. Homopolymer tracts are also abundant, comprising a further 16% of non-coding sequence. The base composition of non-coding SSRs and homopolymer tracts (99.2% A+T content) is even more biased than that of the surrounding sequence, suggesting that either selection or the mechanism of repeat expansion favours (A+T)-rich repeats.

Notably, SSRs are also abundant in protein-coding sequence, occurring on average every 724 bp within exons. We consider these coding SSRs in further detail below, in the context of proteins.

Transposable elements are clustered

The genome is rich in transposable elements^{9,12}. Completion of the sequence confirms the earlier observation that transposable elements of the same type are clustered, suggesting their preferential insertion within similar resident elements. However, none of the elements appears to use a specific sequence as a target for insertion: they insert at random within other elements of the same type. Non-long terminal repeat (LTR) retrotransposons are known to insert next to transfer RNA genes; we find many such instances

(Fig. 2), but again no specific sequences were identified as insertion targets.

tRNAs are numerous and paired by specificity

The sequenced genome encodes 390 tRNAs, a number at the upper end of the eukaryotic spectrum (for example, *Plasmodium falciparum* = 43, *Drosophila melanogaster* = 284, *Homo sapiens* = 496). Allowing for the normal wobble rules in codon–anticodon pairing^{13,14}, every sense codon can be decoded, apart from the rare alanine codon GCG; we infer that the missing tRNA(s) lie in one or more gaps in the sequence. We also find a possible selenocysteine tRNA in the genome, as well as corresponding selenocysteine insertion targets in two predicted proteins (see Supplementary Fig. 3).

Dictyostelium, in common only with *Acanthamoeba castellanii*¹⁵, has been shown to lack certain apparently essential tRNAs in its mitochondrial genome¹⁶. It therefore seems likely that at least some chromosomally encoded tRNAs (those for valine, threonine, asparagine and glycine, as well as one arginine and two serine tRNAs) are imported into mitochondria.

Although the gross distribution of tRNAs is uniform, organization of tRNAs on a finer scale is striking: about 20% occur as pairs or triplets with identical anticodons (and usually 100% sequence identity), separated by <20 kb and often by <5 kb (Fig. 2). There are 41 such groups in the genome; a random distribution would produce few, if any. This pattern is unique among sequenced genomes, and suggests a wave of recent duplications. However, tRNA pairs are found in tandem, converging and diverging orientations with comparable frequencies, suggesting no straightforward duplication mechanism; nor is there usually duplication of extensive flanking sequences. Whether the preference of TRE elements for inserting adjacent to tRNAs is related to the large number and unusual distribution of tRNAs is unclear.

A chromosomal master copy of the extrachromosomal rDNA element

In *Dictyostelium*, ribosomal RNA genes lie on an 88-kb palindromic extrachromosomal element¹⁷, present at ~100 copies per nucleus (Fig. 2). Evidence also exists of chromosomal copies: at least the central 3.2 kb of the element is located¹⁷ on chromosome 4, whereas

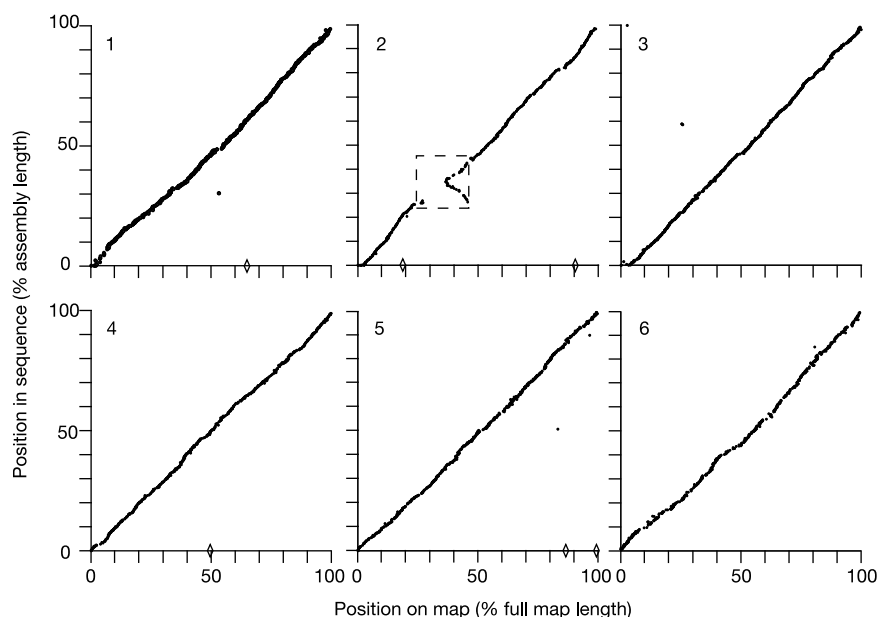


Figure 1 Chromosomal assemblies compared against HAPPY map data. The locations of markers as found in the sequence (y axis) are plotted against their location in HAPPY maps (x axis) for chromosomes 1–6. Markers mapped to one chromosome but found in the

assembled sequence of another are indicated by diamonds on the x axis. The dashed box indicates a large inverted duplication on chromosome 2: markers in this region are shown at one of their two possible map locations but are found at two points in the sequence.

chromosome 2 carries both a partial rDNA sequence and a 5S rRNA pseudogene^{9,18}.

In this study, two unanchored contigs assigned to chromosomes 4 and 5 contained junctions between rDNA sequences and complex repeats—attempts to extend the sequence and integrate these contigs into the assemblies failed owing to the highly repetitive nature of the adjoining sequences. We postulate that these contigs represent the junctions between a 'master copy' of the rDNA and the remainder of chromosome 4 (Fig. 2). One contig contains sequence matching a region of (G+C)-rich repeats near the centre of the palindrome, whereas the other matches sequence near the tip of the palindrome arm, adjacent to the one unclosed gap in the rDNA element sequence¹⁷. This gap is believed to represent a tandem array of short repeats, probably added post-synthetically to the extrachromosomal elements.

The structure of this master copy suggests a mechanism for generating the extrachromosomal copies by a process of transcription, hairpin formation and second-strand synthesis (Fig. 2). This process would account for the complete absence of sequence variation between the two arms of the palindrome.

Centromeres, telomeres and rearrangements

Repeat clusters may serve as centromeres

Centromeres mobilize eukaryotic chromosomes during cell division but vary widely in their structure and organization¹⁹, making them difficult to identify. Each *Dictyostelium* chromosome carries a single cluster of repeats rich in DIRS (*Dictyostelium* intermediate repeat sequence) elements^{20,21} near one end²², and this sole but striking structural consistency suggests that these clusters may serve as centromeres. Although the repetitive nature of the chromosomal termini impeded their assembly, most of the cluster on chromosome 1 was assembled (Fig. 3) and shows a complex pattern of DIRS and related Skipper elements, each preferentially associated with others of the same type. Frequent insertions and partial deletions have created a mosaic with little long-range order.

In *Dictyostelium* cells demonstrating condensed chromosomes characteristic of mitosis, DIRS-element probes hybridize to one end of each chromosome (Supplementary Fig. 4), consistent with the mapping data. DIRS-like elements in other species are more uniformly scattered along the chromosomes²³, suggesting that their restricted distribution in *Dictyostelium* chromosomes is functionally important. Furthermore, the DIRS-containing ends of the chromosomes cluster not only during mitosis, but also during interphase (Supplementary Fig. 4), as has been observed for centromeres in *Schizosaccharomyces pombe*²⁴.

rDNA sequences seem to act as telomeres

No (G+T)-rich telomere-like motifs were identified in the sequence; however, earlier findings²² suggested that the chromosomes terminate in the same (G+A)-rich repeat motif that caps the extrachromosomal rDNA element. We therefore surveyed all shotgun sequence to identify reads containing a junction between complex repetitive elements and rDNA-like sequence. Only 556 such reads were identified, of which 221 could be built into 13 contigs, which we refer to as C/R (complex-repeat/rDNA) junctions.

Of the 13 junctions, two represent known regions lying internally within the chromosomal assemblies. Of the remaining 11, one had twice the sequence coverage of the others, suggesting that it represents two distinct but identical portions of the genome (a possibility supported by the fact that another two of the junctions differed from each other by only two bases). Hence, we infer that the 11 remaining contigs represent 12 distinct junctions between repetitive elements and rDNA-like sequences—potentially one for every chromosomal end.

On the basis of their content of sequence reads from each of the whole-chromosome libraries, we assigned two of the C/R junctions

to each of the chromosomes. Chromosomes 4 and 5 cannot be distinguished in this way, but three junctions, including the one believed to be present as two copies, are assigned to this chromosome pair. The point in the rDNA palindrome that is represented differs from one junction to the next (Supplementary Fig. 5), but several junctions fall at common parts of the palindrome. This may reflect a preference in the mechanism that forms or maintains the junctions, or may result from a homogenizing recombination between them or with other rDNA sequences. Certainly the low frequency of differences between the rDNA components of the junction fragments and the extrachromosomal rDNA element argues for some process that limits or rectifies mutation. At each junction, we see only the rDNA sequence that immediately adjoins the complex repeat, as further assembly is precluded by the multi-copy nature of rDNA. Therefore we cannot tell whether each junctional rDNA sequence extends to the telomere-repeat-carrying tip of the rDNA palindrome sequence, nor whether other sequences lie beyond the rDNA components.

HAPPY mapping of markers derived from six of these C/R junctions confirmed not only the chromosomal assignments that had been made based on the origins of their component sequences, but also their locations at the termini of the mapped regions of the chromosomes. For the other junctions, the absence of unique sequence features precluded such mapping. Taken as a whole, this evidence strongly suggests that rDNA-like elements form part of the telomere structure in *D. discoideum*, and that common mechanisms stabilize both the extrachromosomal rDNA element and the chromosomal termini.

Chromosome 2 duplication

Chromosome 2 of *D. discoideum* AX4 carries a perfect inverted 1.51-

Figure 2 The genome of *Dictyostelium discoideum*. On each of the chromosomal assemblies (numbered 1–6) the diameter of the tube represents coding density (proportion of coding bases summed over both strands; centre-weighted sliding window of 100 kb; scale on right). The coloured bands on the chromosomes represent tRNAs (red), complex repeats (blue), gaps (black) and ribosomal DNA sequences (yellow). G+C content is plotted above each chromosome (centre-weighted sliding window of 100 kb; scale on left). The locations of HAPPY markers are indicated by short green ticks immediately below the distance scale. Immediately beneath each chromosome, the locations (short vertical ticks) of genes known to be upregulated (red), downregulated (blue) or whose level of expression does not change significantly (grey) in the transition from solitary to aggregative existence (expression data from ref. 91) are indicated; coloured horizontal bars below this indicate significant clusters of genes that are preferentially expressed in germinating spores (red), de-differentiating cells (green), pre-spore cells (blue) or in pre-stalk cells (yellow). The translucent 'hourglass' shape on chromosome 2 is centred on a large inverted duplication. The translucent cylinder on chromosome 3 indicates a typical 300-kb region, which is shown in expanded form in inset **a** to illustrate the clustering of identical tRNA genes (red arrows indicate polarity of tRNA genes); a 50-kb section of this region is expanded further in inset panel **b**, revealing the close association of TRE elements (specific family named above) with tRNAs. The translucent yellow disc on chromosome 4 indicates the location of the presumed chromosomal master copy of the rDNA element. In inset panel **c**, the structure of the palindromic extrachromosomal element is shown schematically. (I) Magenta bands indicate rDNA genes; green bands indicate G+C-rich regions; red end caps indicate short repetitive telomere structures; the translucent hoop indicates the central region of asymmetry. (II) Two chromosomal sequence contigs, each carrying an rDNA-like sequence (green or yellow; dotted lines indicate corresponding part of element) flanked by complex repeats (blue). From these contigs, we infer the probable structure (III) of the genomic master copy (grey indicates flanking sequence on chromosome 4). This structure suggests a mechanism for regenerating the extrachromosomal copies by transcription of a single strand (IV), hairpin formation and strand extension (V; broken line indicates synthesis of complementary strand), unfolding of the hairpin and synthesis of a fully complementary strand (VI; broken line indicates synthesis of second strand; telomeric caps added post-synthetically).

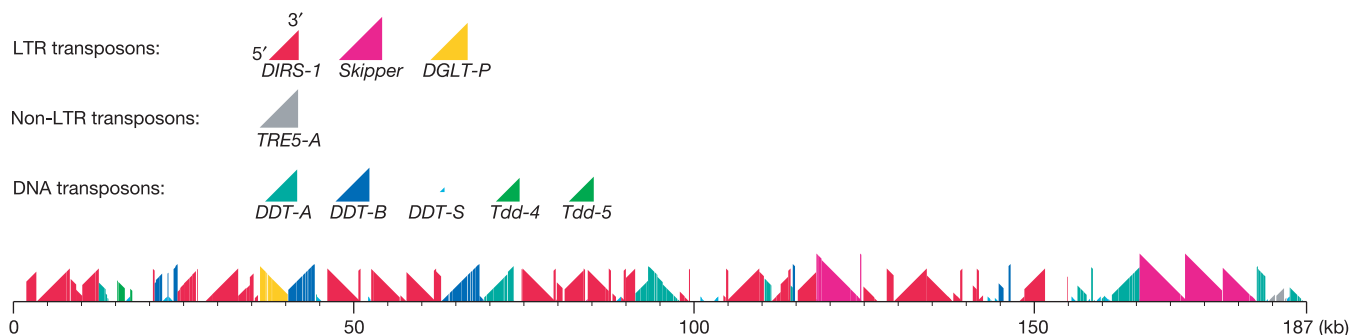


Figure 3 DIRS repeat region of chromosome 1. Complete complex repeat units are represented by coloured triangles whose size corresponds to the sequence length of the repeat unit (see key at top of figure). The bottom-left and top-right corners of each triangle, both at about 3.74 Mb (Fig. 2)—have been implicated in centromeric and telomeric functions, respectively, elsewhere in the genome.

megabase (Mb) duplication (Fig. 2; see also refs 9, 25). This duplication, containing 608 genes, is known²⁵ to be absent from the wild-type isolate NC4 and from one of its direct descendents (AX2), but present in another (AX3); AX4 in turn is derived from AX3. The sequences adjoining the right-hand end of the duplication—a partial copy of a DIRS element (and a partial DDT-A element) and a region identical to part of the rDNA palindrome, both at about 3.74 Mb (Fig. 2)—have been implicated in centromeric and telomeric functions, respectively, elsewhere in the genome.

We propose that this duplication arose from a ‘breakage-fusion-bridge’ cycle as first described in maize²⁶ and since observed in many genomes. The nearby DIRS and rDNA components, in this view, represent abortive attempts to stabilize the halves of the broken chromosome by establishing new telomeres and centromeres, followed by re-fusion of the pieces to create a restored and enlarged chromosome (Supplementary Fig. 6).

Chromosome 2 (the largest of the chromosomes, even discounting the duplication in AX4) may be prone to breakage: in the Bonner isolate of NC4, maintained in vegetative growth for 50 years, chromosome 2 is represented by two smaller fragments²⁷. Comparison with more recent data²² indicates that the break point in NC4-Bonner lies in the same region as the duplication in AX4, suggesting that NC4-Bonner underwent the early stages of this process, but that the chromosome fragments were stabilized and maintained after the initial breakage. Preliminary results (data not shown) from HAPPY mapping also suggest that although wild-type isolates V12M2 and NC4 both lack the duplication seen in AX4, NC4 may carry a duplication of ~300 kb near the opposite end of chromosome 2.

partial repeat units within the first 187 kb of *D. discoideum* chromosome 1 is shown (bottom) by corresponding portions of the triangles; the orientation of the triangles indicates the direction in which each repeat unit lies. The vertical scale (sizes of repeat units) is the same as the horizontal scale (chromosomal distances).

Content and organization of the proteome

Prediction of protein-coding genes (see Methods) was performed on the complete set of chromosomes and floating contigs (Table 2). In assessing the completeness and accuracy of the predictions, we find that of the 957 well-characterized *D. discoideum* genes that are present in the current sequence, 823 (86%) are predicted as transcripts with structures matching the experimentally determined ones. For a further 123 (13%), the predicted transcript differs from the experimentally determined one, about one-half of these differing only in their 5' boundary; the remaining 11 (1%), although present in the sequence, were not predicted as transcripts. Similarly, of the 128,207 qualified ESTs present in the current sequence, 127,097 (99.1%) fall within predicted transcripts. Combining our estimate of sequence coverage (above) with these estimates of the success of gene prediction, we infer that approximately 98% of all *D. discoideum* genes are present in the predicted set.

The level of overprediction, conversely, is harder to estimate: prediction was performed generously to ensure that most true genes were represented. Of the 13,541 predicted proteins, 47.5% are represented by qualified ESTs, reflecting the inevitable bias in EST sampling. Among the shortest predicted proteins, fewer are represented by ESTs (for example, 21% of those of <60 amino acids); this is at least partly due to a higher level of overprediction. On the basis of the simplifying assumption that 50% of all genes coding for proteins of <100 amino acids are mis-predictions, we estimate the true number of genes at roughly 12,500. This number is closer to that seen in multicellular organisms rather than in most unicellular eukaryotes (Table 2). The same relative complexity is seen in the total number of amino acids encoded by the respective genomes; this measure of complexity is less affected by the inclusion

Table 2 Comparison between the predicted protein-coding gene set of *D. discoideum* and those of other organisms

Feature	<i>D. discoideum</i>	<i>P. falciparum</i>	<i>S. cerevisiae</i>	<i>A. thaliana</i>	<i>D. melanogaster</i>	<i>C. elegans</i>	Human
Genome size (Mb)	34	23	13	125	180	103	2,851
Number of genes	12,500*	5,268	5,538	25,498	13,676	19,893	22,287
Gene spacing (kb per gene)	2.5	4.3	2.2	4.9	13.2	5.0	127.9
Mean gene length (bp)	1,756	2,534	1,428	2,036	1,997	2,991	27,000
Mean coding size (amino acids)	518	761	475	437	538	435	509
Genes with introns (%)	69	54	5	79	38	5	85
Mean intron size (bp)	146	179	ND	170	ND	270	3,365
Mean no. of introns (in spliced genes)	1.9	2.6	1.0	5.4	4.0	5.0	8.1
Total amino acids encoded (thousands)	7,021	4,009	2,471	11,143	7,358	9,038	11,333
Codon A + T bias†	86	83	62	57	50	64	41
Mean A + T percentage (exons)	73	76	72	72	45	58	55
Mean A + T percentage (introns)	88	87	51	55	38	71	62
Mean A + T percentage (intergenic)	85	86	51	56	38	72	62

ND, not determined.

* See text. The estimated number of true transcripts for *D. discoideum* is given here for comparability with other species; however, the total predicted gene number of 13,541 is used in calculating the figures below.

† Percentage of all codons used which have A/T at their third base.

of shorter (and hence more dubious) gene predictions. Introns in *Dictyostelium* are few and short, and intergenic regions are small, producing a compact genome of which 62% encodes protein.

Genes are distributed approximately uniformly across the genome (Fig. 2). Although we do not see widespread clustering of genes with coordinated expression patterns (see Methods), we do find statistically significant ($P < 0.01$) clusters of genes expressed predominantly at some developmental stages or in specific cell types (Fig. 2).

(A+T)-richness influences protein composition and codon usage

Codon usage in *Dictyostelium* favours codons of the form NNT or NNA over their NNG or NNC synonyms, the bias being even greater than for the (A+T)-rich *Plasmodium* genome. Comparison of tRNA and codon frequencies (Supplementary Table 2) reveals a similar picture to that in human²⁸ and other eukaryotes, suggesting that the same use is made of 'wobble' and of base modifications (for example, of adenine to inosine in some tRNAs) to expand the effective repertoire of tRNAs.

As in *Plasmodium*²⁹, the extreme (A+T)-richness is reflected not just in the choice of synonymous codons, but also in the amino acid composition of the proteins. Amino acids encoded solely by codons of the form WVN (where W indicates A or T and N indicates any base; these are Asn, Lys, Ile, Tyr and Phe) are much commoner in *Dictyostelium* proteins than in human ones; the reverse is true for those encoded solely by SSN codons (where S indicates C or G; these are Pro, Arg, Ala and Gly).

Geometry reflects phylogeny—duplications in the genome

The predicted gene set of *Dictyostelium* is rich in relatively recently duplicated genes. Of the 13,498 predicted proteins analysed, 3,663 fall into 889 families clustered by BLASTP similarities of $e < 10^{-40}$. Most (538) families contain only two members, but 351 families contain between three and 81 proteins (Supplementary Table 3). Hence, 2,774 (20%) of all predicted proteins have arisen by relatively recent duplication, potentially accounting for much of *Dictyostelium*'s excess gene number compared with typical unicellular eukaryotes.

We tried to infer the mechanisms by which such duplications

arise and propagate in the genome. Where members of a family are clustered on one chromosome, the physical distance between family members often (23 out of 86 families examined) correlates strongly with their evolutionary divergence (see Methods). Where a family is split between different chromosomes, members on the same chromosome are often (23 out of 50 families examined) more related to each other than to members on different chromosomes; the reverse is never observed.

These findings suggest that three processes combine to account for most of the duplications in *Dictyostelium*: tandem duplication, local inversion and interchromosomal exchange. In this model, gene families expand by tandem duplication of either single genes or blocks containing several consecutive genes, as in an earlier model³⁰; inversions within these expanding clusters may reverse local gene order. An elegant illustration of these two processes is provided by a cluster of acetyl-coA synthetases on chromosome 2 (Fig. 4). The third process (exchange of segments between chromosomes) may fragment these clusters at any stage. If such an interchromosomal exchange splits a gene family early in its expansion, then each of the two resulting subfamilies has a long subsequent period of evolution independent of the other, so similarities will be greatest between genes on the same chromosome. If, conversely, the split occurs later, then all family members, whether on the same chromosome or on different chromosomes, will tend to resemble each other equally closely. We cannot exclude the possibility of duplication occasionally creating a second copy of a gene, or group of genes, directly on a different chromosome from the first. However, all instances that we have examined can be accounted for without such intermolecular duplication.

Amino acid repeats

Tandem repeats of trinucleotides (and of motifs of 6, 9, 12, and so on, bases) are unusually abundant in *Dictyostelium* exons and naturally correspond to repeated sequences of amino acids. However, at the protein level the situation is even more extreme: there are many further amino acid repeats that use different synonymous codons, and so do not arise from perfect nucleotide repeats. Among the predicted proteins, there are 9,582 SSRs of amino acids (homopolymers of length ≥ 10 , or ≥ 5 consecutive repeats of a motif of two

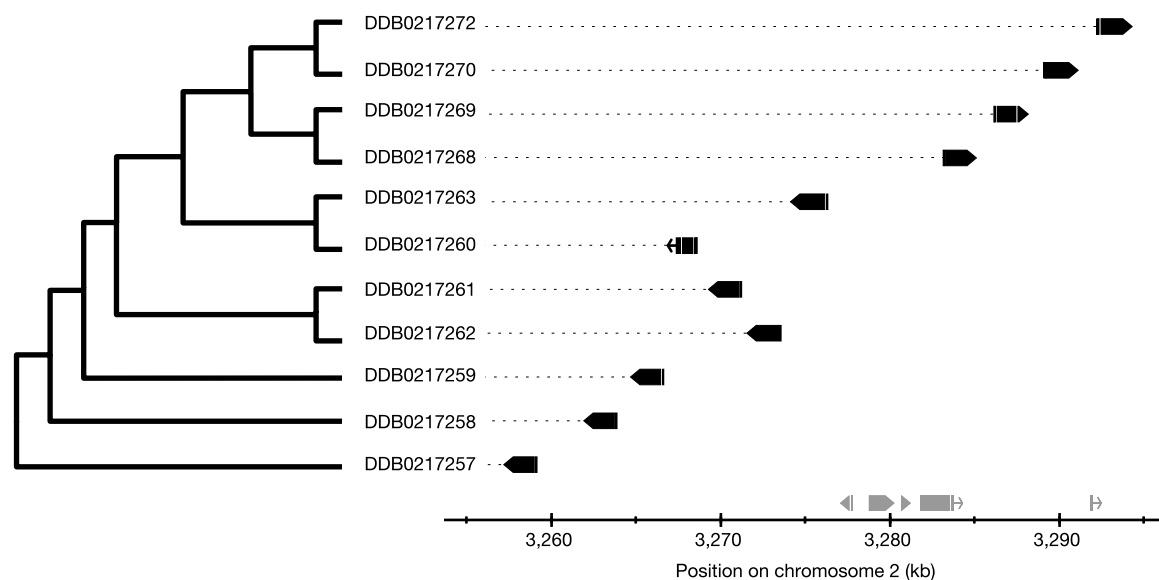


Figure 4 Phylogeny of gene family members compared to their physical order. The optimally parsimonious phylogenetic tree of 11 acetyl-CoA synthase genes, computed using the PHYLIP module 'Protpars' (<http://evolution.gs.washington.edu/phylip/doc/protpars.html>), is shown to the left; dictyBase identification numbers are shown at the end of each branch. The graph (right) indicates the arrangement on chromosome 2 of these

genes (solid black boxes; gaps indicate introns; pointed ends indicate direction of transcription). Chromosomal distance scale is given along the bottom and other unrelated genes in the same region are indicated in grey above the x axis. The correspondence between phylogeny and physical order implies that the cluster has arisen by a series of segmental tandem duplications and local inversions in parallel with sequence divergence.

or more amino acids). Of these, the most striking are polyasparagine and polyglutamine tracts of ≥ 20 residues, present in 2,091 of the predicted proteins. Also abundant are low-complexity regions such as QLQLQQQQQLQLQQ: there are 2,379 tracts of ≥ 15 residues composed of only two different amino acids. In total, repeats or simple-sequence tracts of amino acids (even by these conservative definitions) occur in 34% of predicted proteins and encode 3.3% of all amino acids.

It seems likely that these repeats have arisen through nucleotide expansion, but have been selected at the protein level. Evidence for selection at the protein level is that any given trinucleotide repeat occurs predominantly in only one of the three reading frames. For example, the repeat ...ACAACAACAACA... is usually translated as polyglutamine ([CAA] n) rather than polythreonine ([ACA] n) or polyasparagine ([AAC] n). Further evidence comes from the many trinucleotide repeats that have apparently mutated to produce only synonymous codons (for example, ...GATGACGATGATGAC..., translated as polyaspartate). Moreover, the distribution of repeats and simple-sequence tracts is nonrandom: most proteins either have no such features (66% of proteins) or have two or more (18% of proteins), suggesting that they are tolerated only in certain types of protein. The polyasparagine- and polyglutamine-containing proteins appear to be over-represented in protein kinases, lipid kinases, transcription factors, RNA helicases and messenger RNA binding proteins such as spliceosome components (Supplementary Fig. 9). Protein kinases and transcription factors are also over-represented in the polyasparagine- and polyglutamine-containing proteins of *Saccharomyces cerevisiae*, so it is possible that these homopolymers serve some functional role in these protein classes. A more detailed analysis of amino acid homopolymers is given in Supplementary Tables 4–6 and Supplementary Figs 7–10.

Phylogeny, evolution and comparative proteomics

The organisms that diverged from the last common ancestor of all eukaryotes followed different evolutionary paths, but all retained the basic properties of eukaryotic cells. Their genomes have been sculpted by chromosomal deletions and duplications that led to lineage-specific gene family expansions, reductions and losses, as

well as genes with new functions^{31,32}. Our analysis of *Dictyostelium*'s proteome shows that similar mechanisms have shaped its genome, augmented by horizontal gene transfer from bacterial species.

Phylogeny of eukaryotes based on complete proteomes

Using morphological criteria, early workers were unsure whether to classify Dictyostelids as fungi or protozoa³³. Molecular methods indicated that they were amoebozoia and also suggested that *Dictyostelium* diverged from the line leading to animals at about the same time as plants^{34,35}. A study of more than 100 proteins suggested that *Dictyostelium* diverged after the plant–animal split, but before the divergence of the fungi³⁶. The recent finding of a gene fusion encoding three pyrimidine biosynthetic enzymes, shared only by *Dictyostelium*, fungi and Metazoa, indicates that the amoebozoia are a true sister group of the fungi and Metazoa³⁷.

To examine the phylogeny of *Dictyostelium* on a genomic scale, we applied an improved method for predicting orthologous protein clusters to complete eukaryotic proteomes³⁸ (for details, see Supplementary Information). The data were used to construct a phylogenetic tree that confirms the divergence of *Dictyostelium* along the branch leading to the Metazoa soon after the plant–animal split (Fig. 5). Despite the earlier divergence of *Dictyostelium*, many of its proteins are more similar to human orthologues than are those of *S. cerevisiae*, probably due to higher rates of evolutionary change along the fungal lineage. Whether the greater similarity between amoebozoia and Metazoa proteins translates into a generally higher degree of functional conservation between them compared to the fungi remains to be seen.

Proteins shared by *Dictyostelium* and major organism groups

To examine shared functions, we identified eukaryote-specific Superfamily and Pfam protein domains, and sorted them according to their presence or absence within 12 completely sequenced genomes to arrive at their distribution among the major organismal groups (see Supplementary Tables 7–10 and Supplementary Fig. 11). Plants, Metazoa, fungi and *Dictyostelium* all share 32% of the eukaryotic Pfam domains (Fig. 6). The protein domains present

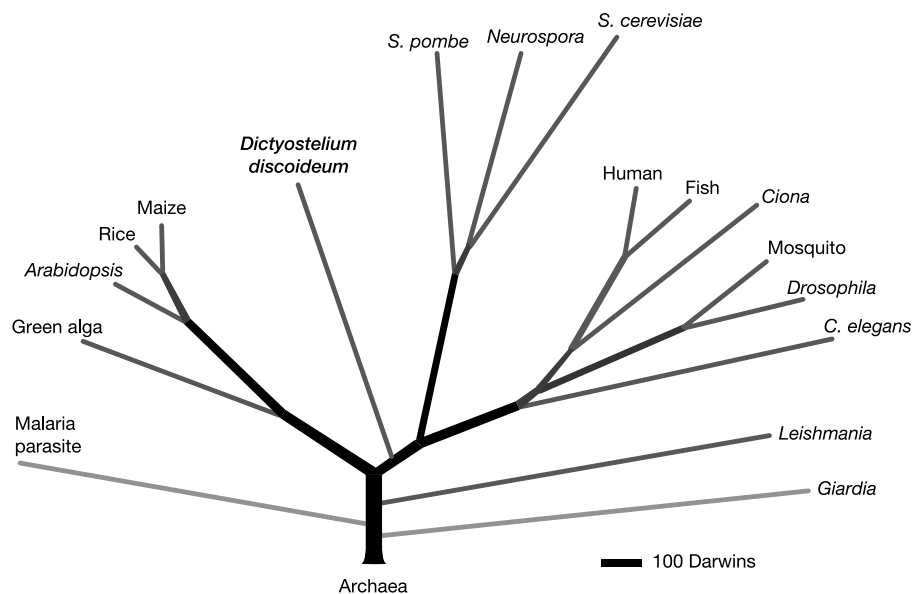


Figure 5 Proteome-based eukaryotic phylogeny. The phylogenetic tree was reconstructed from a database of 5,279 orthologous protein clusters drawn from the proteomes of the 17 eukaryotes shown, and was rooted on 159 protein clusters that had representatives from six archaeobacterial proteomes. Tree construction, the database of protein clusters and a model of protein divergence used for maximum likelihood

estimation are described in Supplementary Information. The relative lengths of the branches are given as Darwins (where 1 Darwin = 1/2,000 of the divergence between *S. cerevisiae* and humans). Species that are not specified are *Plasmodium falciparum* (malaria parasite), *Chlamydomonas reinhardtii* (green alga), *Oryza sativa* (rice), *Zea mays* (maize), *Takifugu rubripes* (fish) and *Anopheles gambiae* (mosquito).

in *Dictyostelium*, Metazoa and fungi, but absent in plants, are interesting because they probably arose soon after plants diverged and before *Dictyostelium* diverged from the line leading to animals. The major classes of domains in this group of proteins include those involved in small and large G-protein signalling (for example, RGS proteins), cell cycle control and other domains involved in signalling (Supplementary Tables 8 and 9). It also appears that glycogen storage and usage arose as a metabolic strategy soon after the plant–animal divergence, because glycogen synthetase seems to have appeared in this evolutionary interval.

Particularly notable are the cases where otherwise ubiquitous domains appear to be completely absent in one group or another. For instance, *Dictyostelium* seems to have lost the genes that encode collagen domains, the circadian rhythm control protein timeless and basic helix–loop–helix transcription factors (Supplementary Table 7). Metazoa, on the other hand, appear to have lost receptor histidine kinases that are common in bacteria, plants and fungi, whereas *Dictyostelium* has retained and expanded its complement to 14 members³⁹.

Orthologues of human disease genes

An important motivation for sequencing the *Dictyostelium* genome was to aid the discovery of proteins that would facilitate studies of orthologues in human, with possible implications for human health. Although orthologues of human genes implicated in disease are of course present in many species, *Dictyostelium* provides a potentially valuable vehicle for studying their functions in a system that is experimentally tractable and intermediate in complexity between the yeasts and the higher multicellular eukaryotes. To assess the usefulness of *Dictyostelium* for investigating the functions of genes related to human disease we used the protein sequences of 287 confirmed human disease genes as queries and carried out a systematic search for putative orthologues in the *Dictyostelium* proteome⁴⁰. At a stringent threshold value of $e \leq 10^{-40}$, we identified 64 such proteins. Of these, 33 were similar in length to the human protein and had similarity extending over >70% of the two proteins (Table 3). The number of *Dictyostelium* orthologues of human disease genes is lower than in *D. melanogaster* or *Caenorhabditis elegans* but higher than in *S. cerevisiae* or *S. pombe*. Of the 33 putative orthologues of confirmed human disease genes in *Dictyostelium*, five are absent in both *S. cerevisiae* and *S. pombe* (e -value $\leq 10^{-30}$), a further four are absent from *S. cerevisiae* and two are not found in *S. pombe*.

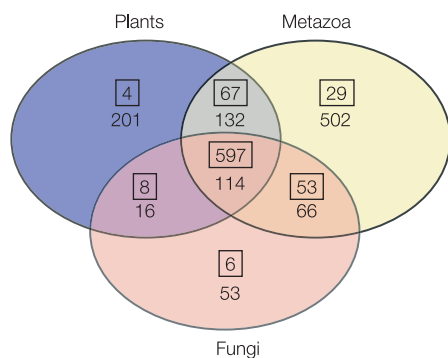


Figure 6 Distribution of Pfam domains among eukaryotes. The number of eukaryote-specific Pfam domains present in each group of eukaryotic organisms is shown. The boxed numbers are the domains that are present in *Dictyostelium* and the other numbers are those domains that are absent from *Dictyostelium*. The animals are *H. sapiens*, *T. rubripes*, *C. elegans*, *D. melanogaster*; the fungi are *N. crassa*, *Aspergillus nidulans*, *S. pombe* and *S. cerevisiae*; and the plants are *Arabidopsis thaliana*, *O. sativa* and *C. reinhardtii*. A complete listing of the domains can be found in the Supplementary Information.

Horizontal gene transfer

The acquisition of genes by horizontal transfer from one species to another (HGT) has become increasingly recognized as a mechanism of genome evolution^{41–43}. We identified 18 potential instances of HGTs, by screening *Dictyostelium* protein domains that are similar to bacteria-specific Pfam domains and have phyletic relationships consistent with HGT (see Supplementary Information). The transferred domains appear to have replaced functions, added new functions or evolved into new functions (Table 4). The *thy1* gene, which encodes an alternative form of thymidylate synthase (ThyX), appears to have replaced the endogenous gene, as the conventional thymidylate synthase (ThyA) is not present⁴⁴. Other HGT domains also have established functions, which are presumably retained and give *Dictyostelium* the ability to degrade bacterial cell walls (dipeptidase), scavenge iron (siderophore), or resist the toxic effects of tellurite in the soil (terD). Still other horizontally transferred domains have become embedded within *Dictyostelium* genes that encode larger proteins. An example of this is the Cna B domain that is found within four large predicted proteins, one of which, colossin A, is predicted to be 1.2 MDa (Supplementary Fig. 12).

Dictyostelium ecology

Dictyostelium faces many complex ecological challenges in the soil. Amoebae, fungi and bacteria compete for limited resources in the soil while defending themselves against predation and toxins. For instance, the nematode *C. elegans* is a competitor for bacterial food

Table 3 *Dictyostelium* genes related to human disease genes

Disease category*	SwissProt†	dictyBase ID‡
Cancer		
Colon cancer (<i>MSH2</i>)	MSH2_HUMAN	DDB0202539
Colon cancer (<i>MLH1</i>)	MLH1_HUMAN	DDB0187465
Colon cancer (<i>MSH3</i>)	MSH3_HUMAN	DDB0204604
Colon cancer (<i>PMS2</i>)	PMS2_HUMAN	DDB0185791
Xeroderma pigmentosum (<i>ERCC3</i>)	XPB_HUMAN	DDB0206281
Xeroderma pigmentosum (<i>XPD</i>)	XPD_HUMAN	DDB0189539
Oncogene (<i>AKT2</i>)	AKT2_HUMAN	DDB0189970
Oncogene (<i>RAS</i>)	RASH_HUMAN	DDB0191937
Cyclin-dependent kinase 4 (<i>CDK4</i>)	CDK4_HUMAN	DDB0188077
Neurological		
Lowe oculocerebrorenal (<i>OCRL</i>)	OCRL_HUMAN	DDB0189888
Miller–Dieker lissencephaly (<i>PAF</i>)	LIS1_HUMAN	DDB0219335
Adrenoleukodystrophy (<i>ABCD1</i>)	ALD_HUMAN (P)	DDB0219834
Angelman (<i>UBE3A</i>)	UE3A_HUMAN	DDB0188760
Ceroid lipofuscinosis (<i>CLN2</i>)	TPP1_HUMAN (C, P)	DDB0190668
Tay–Sachs (<i>HEXA</i>)	HEXA_HUMAN (C, P)	DDB0187255
Ceroid lipofuscinosis (<i>PPT</i>)	PPT1_HUMAN (C)	DDB0186550
Thomsen myotonia congenita (<i>CLCN1</i>)	CLC1_HUMAN	DDB0191805
Choroideremia (<i>CHM</i>)	RAE1_HUMAN	DDB0206402
Amyotrophic lateral sclerosis (<i>SOD1</i>)	SODC_HUMAN	DDB0188850
Parkinson's (<i>UCHL1</i>)	UCL1_HUMAN (C, P)	DDB0205083
Cardiovascular		
Hypertrophic cardiomyopathy	MYH7_HUMAN	DDB0186963
Renal		
Renal tubular acidosis (<i>ATP6B1</i>)	VAB1_HUMAN	DDB0169211
Hyperoxaluria (<i>AGXT</i>)	SPYA_HUMAN (C, P)	DDB0188646
Metabolic/endocrine		
Niemann–Pick type C (<i>NPC1</i>)	NPC1_HUMAN (P)	DDB0191057
Hyperinsulinism (<i>ABCC8</i>)	ACC8_HUMAN	DDB0187670
McCune–Albright (<i>GNAS1</i>)	GBAS_HUMAN	DDB0185461
Pendred (<i>PDS</i>)	PEND_HUMAN (C)	DDB0202939
Haematological/immune		
G6PD deficiency (<i>G6PD</i>)	G6PD_HUMAN	DDB0168147
Chronic granulomatous (<i>CYBB</i>)	C24B_HUMAN (C, P)	DDB0188527
Malformation		
Diastrophic dysplasia (<i>SLC26A2</i>)	DTD_HUMAN (C)	DDB0202939
Other		
Cystic fibrosis (<i>ABCC7</i>)	CFTR_HUMAN	DDB0186232
Darier–White (<i>SERCA</i>)	ATA2_HUMAN	DDB0169159
Congenital chloride diarrhoea (<i>DRA</i>)	DRA_HUMAN (C)	DDB0202939

*From a list of 287 confirmed human disease protein sequences⁴⁰. Those listed match a predicted *Dictyostelium* protein with a BLASTP probability of $e \leq 10^{-40}$, are similar in length ($\pm 25\%$ in comparison to the *Dictyostelium* protein) and both proteins align over more than 70% of their respective lengths.

†SwissProt identifiers for the human proteins. Letters in brackets indicate that the protein has no homologue (BLASTP probability of $e \leq 1.0 \times 10^{-30}$) in *S. cerevisiae* (C) or *S. pombe* (P).

‡The best match to the human gene is listed by its dictyBase identification number. Matches with a BLASTP probability of $e \leq 10^{-100}$ are indicated in bold.

and a predator of *Dictyostelium* amoebae, but also a potential dispersal agent for *Dictyostelium* spores⁴⁵. *Dictyostelium* has expanded its repertoire of several protein classes that are probably crucial for such interspecies interactions and for survival and motility in this complex ecosystem.

Polyketide synthases

A small number of natural products have already been identified from *Dictyostelium*, but the gene content suggests that it is a prolific producer of such molecules. Some of them may act as signals during development, such as the dichlorohexanophenone DIF-1, but others are likely to mediate currently unknown ecological interactions⁴⁶. Many antibiotics and secondary metabolites destined for export are produced by polyketide synthases, modular proteins of around 3,000 amino acids⁴⁷. We identified 43 putative polyketide synthases in *Dictyostelium* (see Supplementary Information). By contrast, *S. cerevisiae* completely lacks polyketide synthases and *Neurospora crassa* has only seven. Furthermore, two of the *Dictyostelium* proteins have an additional chalcone synthase domain, representing a type of polyketide synthase most typical of higher plants and found to be exclusively shared by *Dictyostelium*, fungi and plants. In addition to polyketide synthases, the predicted proteome has chlorinating and dechlorinating enzymes as well as O-methyl transferases, which could increase the diversity of natural products made. Thus, *Dictyostelium* appears to have a large secondary metabolism, which warrants further investigation.

ABC transporters

ATP-binding cassette (ABC) transporters are prevalent in the proteomes of soil microorganisms and are thought to provide resistance to xenobiotics through their ability to translocate small-molecule substrates across membranes against a substantial concentration gradient^{48–51}. There are 66 ABC transporters encoded by the genome, which can be classified according to the subfamilies defined in humans (ABCA, ABCB, ABCC, ABCD, ABCE, ABCF and ABCG) based on domain arrangement and signature sequences⁵². At least 20 of them are expressed during growth and are probably involved in detoxification and the export of endogenous secondary metabolites.

Cellulose degradation

Many of the predicted cellulose-degrading enzymes in the proteome

(see Supplementary Information) that have secretion signals are expressed in growing cells that do not produce cellulose⁵³. The proteome also contains one xylanase enzyme that can degrade the xylan polymers that are often found associated with the cellulose of higher plants. Perhaps *Dictyostelium* uses these enzymes to degrade plant tissue into particles that are then taken up by cells. These enzymes may also aid in the breakdown of cellulose-containing microorganisms upon which *Dictyostelium* feeds. Alternatively, these enzymes may promote the growth of bacteria that can serve as food, because *Dictyostelium*'s habitat also contains cellulose-degrading bacteria.

Specializations for cell motility

During both growth and development, *Dictyostelium* amoebae display motility that is characteristic of human leukocytes⁵⁴. As a consequence, studies of *Dictyostelium* have contributed significantly to cytoskeleton research⁵⁵. *Dictyostelium*'s survival depends on an ability to efficiently sense, track and consume soil bacteria using sophisticated systems for chemotaxis and phagocytosis. Its multicellular development depends on chemotactic aggregation of individual amoebae and the coordinated movement of thousands of cells during fruiting body morphogenesis. The proteome reveals an astonishing assortment of proteins that are used for robust, dynamic control of the cytoskeleton during these processes. As suggested by functional parallels to human cells, these proteins are most similar to metazoan proteins in their variety and domain arrangements (Fig. 7; see also Supplementary Table 11). Surprisingly, although the actin cytoskeleton has been studied for over 25 years, 71 putative actin-binding proteins apparently escaped classical methods of discovery. For example, actobindins had not been previously recognized in *Dictyostelium*. Curiously, the actin depolymerization factor (ADF) and calponin homology (CH) domain proteins appear to have diversified by domain shuffling, a substantial fraction having domain combinations unique to *Dictyostelium* (Supplementary Table 12 and Supplementary Fig. 13). In addition to 30 actin genes, there are also orthologues of all actin-related protein (ARP) classes present in mammals, as well as three founding members of a new class (Supplementary Fig. 14).

Cytoskeletal remodelling during chemotaxis and phagocytosis is regulated by a considerable number of upstream signalling components. Of the 18 Rho family GTPases in *Dictyostelium*, some are

Table 4 Candidate horizontal gene transfers from bacteria

Function*	Pfam†	Number of proteins‡	DictyBase ID§	Length (aa)	Region matched¶	e-value#
Aromatic amino acid lyase	Beta_elim_lyase	2	DDB0204031	170	4–170	3.2×10^{-65}
Biotin metabolism	BioY	1	DDB0184375	338	145–299	5.8×10^{-20}
Unknown	Cna_B	4	DDB0184530	11,103	Multiple††	1.1×10^{-10}
Peroxidase	Dyp_peroxidase	1	DDB0168077	306	3–303	1.4×10^{-82}
Insecticide	Endotoxin_N	2	DDB0188332	628	38–210	1.2×10^{-32}
Isopentenyl transferase	IPT	1	DDB0169077	283	1–63	5.1×10^{-12}
Siderophore	lucA_lucC	2	DDB0219918	739	183–350	2.3×10^{-18}
Osmoregulation	OsmC	2	DDB0190102	156	16–156	9.8×10^{-22}
Dipeptidase/β-lactamase	Peptidase M15	1	DDB0205124	897	68–406; 711–879	3.4×10^{-16}
Dipeptidase/β-lactamase	Peptidase S13	1	DDB0168572	522	337–495	4.2×10^{-25}
Polyphosphate synthesis	PP_kinase	1	DDB0192001	1,053	372–1045	1.6×10^{-234}
Tellurite resistance	TerD	2	DDB0169240	287	152–279	2.1×10^{-67}
Thymidylate synthesis	Thy1	1	DDB0214905	303	38–254	9.9×10^{-117}
Unknown	DUF84	1	DDB0203145	179	5–175	1.6×10^{-20}
Unknown	DUF885	2	DDB0205394	689	318–685	1.5×10^{-124}
(Prespore protein 3B)	DUF1121	3	DDB0169184	226	1–226	8.7×10^{-134}
Unknown	DUF1289	1	DDB0204782	88	29–85	3.3×10^{-15}
Unknown	DUF1294	1	DDB0186703	155	2–73	8.9×10^{-18}

* Confirmed or proposed function of the prokaryotic orthologue is given. For domains without function information, information on any *Dictyostelium* protein in the set is given in parentheses.

† The Pfam domain designation (<http://www.sanger.ac.uk/Software/Pfam/>).

‡ The number of gene models in which the domain appears. Bold numbers indicate gene sets where there are pairs of genes that map within 10 kb of each other.

§ The gene identification number for the example given in the rest of the table (release v2.0 at <http://www.dictybase.org/>).

|| Number of amino acid (aa) residues in the predicted *Dictyostelium* protein containing the domain.

¶ The region of the *Dictyostelium* protein that matched the prokaryotic domain. The amino acid sequence identity between this region and the most highly related prokaryotic protein was 21–52%.

The e-value for the domain against the Pfam model library used to identify it (see Supplementary Information).

†† The protein colossin A consists of an array of 91 partial Cna_B domains within 18 larger repeats, and the e-value corresponds to one domain.

Class	Protein	No.	Module	Occurrence			
				D	M	F	P
G-actin binding	Profilin	3	PRO				
	Actobindin-like	3	WH2				
	CAP	1	WH2				
	WH2-containing	5	WH2				
	Twinfilin-like	1	ADP				
Capping and/or severing	Cap32/34 (Aginactin)	2	CAP				
	Cofilin	6	ADF				
	Severin	1	GEL				
	GRP125	1	GEL				
	Gelsolin-like	2	GEL				
Actin capping and nucleation	Arp2/3 complex	7	ACT				
	Scar	1	WH2				
	WASP	1	WH2				
	WASP-related	2	WH2				
	VASP	1	EVH				
Actin cross-linking	Formin	10	FH2				
	ABP34	1					
	eEF1- α (ABP50)	2	EF1- α				
	eEF1- β	3					
	Dynactin	1					
	Fimbrin	1	CH				
	Fimbrin type ABD-containing	5	CH				
	Filamin (gelation factor)	1	CH				
	α -actinin	1	CH				
	Cortexillin	2	CH				
	α -actinin type ABD-containing	3	CH				
	Protovillin (Cap100)	1	GEL				
	Villin-related	1	GEL				
	Flightless/Villin-related	1	GEL				
	Villin	1	GEL				
Lateral actin binding	Kelch-related	1	KELCH				
	Smoothelin-related	1	CH				
	GAS2-related	1	CH				
	CH-containing	19	CH				
	VHP-containing	3	VHP				
	Coronin	1					
	Coronin-like	1					
	Aip	1					
	Coactosin	1	ADF				
	Coactosin-related	3	ADF				
Membrane-associated	Abp1	1	ADF				
	Glia maturation factor-related	1	ADF				
	UIM domain-containing	3					
	Interaptin	1	CH				
	Ponticulin	2					
	Ponticulin-related	2					
	Comitin	1					
	Comitin-related	1					
	Hisactophilin	3	TRE				
	Talin A (filopodin)	1	TAL				
Motors	Talin B	1	TAL				
	SLA-2-like	1	TAL				
	Annexin	2					
	Vinculin/ α -catenin-related	2					
	Conventional myosin	1	MYO				
	Unconventional myosins	12	MYO				
Total 138 (71 new)							

Figure 7 Microfilament system proteins. Proteins with probable interactions with the actin cytoskeleton are tabulated by their documented or predicted functions. Coloured boxes indicate the presence of a protein related to the *Dictyostelium* (D) protein in Metazoa (M), fungi (F) or plants (P). *Dictyostelium*-specific proteins have no recognizable relatives or differ from relatives due to extensions or unusual domain compositions. For details see Supplementary Information. Actin-binding modules: ACT, actin fold; ADF, actin depolymerization factor/cofilin-like domain; CAP, capping protein fold; CH, calponin homology domain; EVH, Ena/VASP homology domain 2; FH2, formin homology 2 domain; GEL, gelsolin repeat domain; KELCH, Kelch repeat domain; MYO, myosin motor domain; PRO, profilin fold; TAL, the I/LWEQ actin-binding domain of talin and related proteins; TRE, trefoil domain; VHP, villin head piece; WH2, Wiskott Aldrich syndrome homology region 2.

clear Rac orthologues and one belongs to the RhoBTB subfamily⁵⁶. However, the Cdc42 and Rho subfamilies characteristic of Metazoa and fungi are absent, as are the Rho subfamily effector proteins. The activities of these GTPases are regulated by two members of the RhoGDI family, by components of ELMO1–DOCK180 complexes and by a large number of proteins carrying RhoGEF and RhoGAP domains (>40 of each), most of which show domain compositions not found in other organisms. Remarkably, *Dictyostelium* appears to be the only lower eukaryote that possesses class I phosphatidylinositol-3-OH kinases, which are at the crossroad of several critical signalling pathways (for details of the regulators and their effectors, see Supplementary Table 13)⁵⁷. The diverse array of these regulators and the discovery of many additional actin-binding proteins suggest that there are many aspects of cytoskeletal regulation that have yet to be explored.

Multicellularity and development

The evolution of multicellularity was arguably as significant as the origin of the eukaryotic cell in enabling the diversification of life. The common unicellular ancestor of the crown group of organisms must have possessed the basic machinery to regulate nutrient uptake, metabolism, cellular defence and reproduction, and it is likely that these mechanisms were adapted to integrate the functions of cells in multicellular organisms. *Dictyostelium* achieved multicellularity through a different evolutionary route compared with plants and animals, yet the ancestors of these respective groups probably started with the same endowment of genes and faced the same problem of achieving cell specialization and tissue organization.

When starved, *Dictyostelium* develops as a true multicellular organism, organizing distinct tissues within a motile slug and producing a fruiting body comprised of a cellular, cellulosic stalk supporting a bolus of spores⁴. Thus, *Dictyostelium* has evolved differentiated cell types and the ability to regulate their proportions and morphogenesis. A broad survey of proteins required for multicellular development shows that *Dictyostelium* has retained cell adhesion and signalling modules normally associated exclusively with animals, whereas the structural elements of the fruiting body and terminally differentiated cells clearly derive from the control of cellulose deposition and metabolism now associated with plants. The *Dictyostelium* genome offers a first glimpse of how multicellularity evolved in the amoebozoan lineage. In the following sections, we consider some of the systems that are particularly relevant to cellular differentiation and integration in a multicellular organism.

Signal transduction through G-protein-coupled receptors

The needs of multicellular development add greatly to those of chemotaxis in demanding dynamically controlled and highly selective signalling systems. G-protein-coupled cell surface receptors (GPCRs) form the basis of such systems in many species, allowing the detection of a variety of environmental and intra-organismal signals such as light, Ca²⁺, odorants, nucleotides and peptides. They are subdivided into six families, which, despite their conserved secondary domain structure, do not share significant sequence similarity⁵⁸. Until recently, in *Dictyostelium* only the seven CAR/CRL (cAMP receptor/ cAMP receptor-like) family GPCRs had been examined in detail^{59,60}. Surprisingly, a detailed search uncovered 48 additional putative GPCRs of which 43 can be grouped into the secretin (family 2), metabotropic glutamate/GABA_B (family 3) and the frizzled/smoothed (family 5) families of receptors (Fig. 8; see also Supplementary Information). The presence of family 2, 3 and 5 receptors in *Dictyostelium* was surprising because they had been thought to be specific to animals. Their occurrence in *Dictyostelium* suggests that they arose before the divergence of the animals and fungi and were later lost in fungi, and that the radiation of GPCRs pre-dates the divergence of the animals and fungi. The secretin

family is particularly interesting because these proteins were thought to be of relatively recent origin, appearing closer to the time of the divergence of animals⁶¹. The putative *Dictyostelium* secretin GPCR does not contain the characteristic GPCR proteolytic site, but its transmembrane domains are clearly more closely related to secretin GPCRs than to other families (Fig. 8). Many downstream signalling components that transduce GPCR signals could also be recognized in the proteome, including heterotrimeric G-protein subunits (fourteen G α , two G β and one G γ proteins) and seven regulators of G-protein signalling (RGS) that share highest similarity with the R4 subfamily of mammalian RGS proteins.

SH2 domain signalling

In animals, SH2 domains act as regulatory modules of proteins in intracellular signalling cascades, interacting with phosphotyrosine-containing peptides in a sequence-specific manner. *Dictyostelium* is the only organism, outside of the animal kingdom, where SH2 domain phosphotyrosine signalling has been shown to occur⁶². What has been lacking in *Dictyostelium* is evidence of the other components of such signalling pathways; that is, equivalents of the metazoan SH2-domain-containing receptors, adaptors and targeting proteins. Three newly predicted proteins are strong candidates for these roles (Supplementary Fig. 15). One of them, CblA, is highly related to the metazoan Cbl proto-oncogene product. This is entirely unexpected because it is the first time that a Cbl homologue has been observed outside the animal kingdom. The Cbl protein is a 'RING finger' ubiquitin-protein ligase that recognizes activated

receptor tyrosine kinases and various molecular adaptors⁶³. Remarkably, the Cbl SH2 domain went unrecognized in the protein sequence, but it was revealed when the crystal structure of the protein was determined⁶⁴. Thus, although SH2 domain proteins are less prevalent in *Dictyostelium*, there is the potential for the kind of complex interactions that typify metazoan SH2 signalling pathways.

ABC transporter signalling

Dictyostelium, like other organisms, has adapted ABC transporters to control various developmental signalling events. Several ABC transporters (TagA, TagB and TagC) are used for peptide-based signalling, similar to that previously observed for mating in *S. cerevisiae* and antigen presentation in human T cells^{65–67}. The novel domain arrangement of the Tag proteins—a serine protease domain fused to a single transporter domain—suggests that they have been selected for improved efficiency in signal production. Additional ABC transporters are needed for cell fate determination in *Dictyostelium*, suggesting that this ubiquitous protein family may be used in similar developmental contexts within many different species⁶⁸.

Kinases and transcription factors

Much cellular signal transduction involves the regulation of protein function through phosphorylation by protein kinases, often leading to the reprogramming of gene transcription in response to extracellular signals. The *Dictyostelium* proteome contains 295 predicted protein kinases, representing as wide a spectrum of kinase families

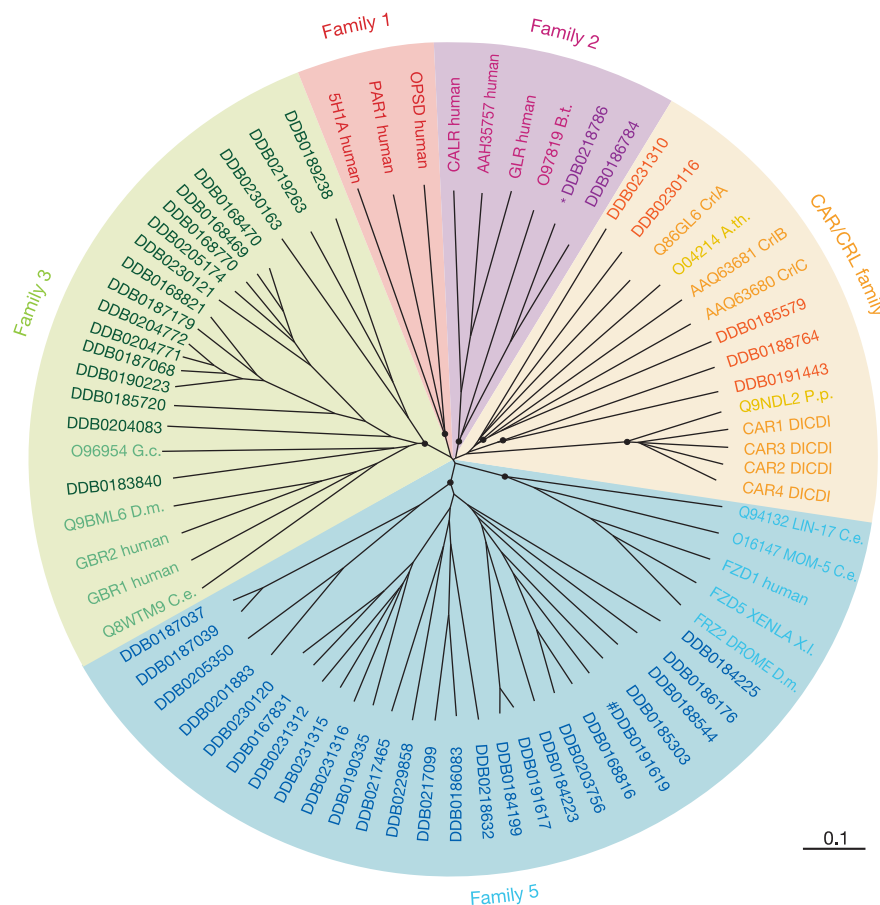


Figure 8 The G-protein-coupled receptors. A CLUSTALX alignment of the sequences encompassing the seven transmembrane domains of all *Dictyostelium* GPCRs, and selected GPCRs from other organisms, was used to create an unrooted dendrogram with the TreeView program. A black circle marks the innermost node of each branch supported by >60% bootstraps. The hash symbol indicates that this gene model has to be split, and

the asterisk indicates a putative pseudogene. DictyBase identifiers (DDB) were used for the newly discovered *Dictyostelium* receptors and SwissProt identifiers for all other receptors. A.th., *A. thaliana*; B.t., *Bos taurus*; CAR/CRL, CAMP receptor/CAMP receptor-like; C.e., *C. elegans*; DICD1, *D. discoideum*; D.m., *D. melanogaster*; G.c., *Geodia cydonium*; P.p., *Polysphondylium pallidum*; X.l., *Xenopus laevis*.

as that observed in Metazoa (Supplementary Tables 14–16 and Supplementary Fig. 16). Given the presence of SH2-domain-based signalling it was surprising that no receptor tyrosine kinases could be recognized in the genome. However, *Dictyostelium* has a number of other receptor kinases, such as the histidine kinases and a group of eight novel putative receptor serine/threonine kinases, which are involved in nutrient and starvation sensing⁶⁹. Most of the ubiquitous families of transcription factors are represented in *Dictyostelium*, with the notable exception of the otherwise ubiquitous basic helix–loop–helix proteins (Supplementary Table 17 and Supplementary Fig. 17). Compared with other eukaryotes, *Dictyostelium* appears to have fewer transcription factors relative to the total number of genes, suggesting that many transcription factors have yet to be defined, or that the activities of a smaller repertoire of factors are combined and controlled to achieve complex regulation (Supplementary Table 18 and Supplementary Fig. 18).

Cell adhesion

Throughout *Dictyostelium* development, cells must modulate their adhesiveness to the substrate, to the extracellular matrix and to other cells in order to create tissues and carry out morphogenesis. To accomplish this, *Dictyostelium* uses a surprising number of components that have been normally only associated with animals. For example, disintegrin proteins regulate cell adhesiveness and differentiation in a number of Metazoa, and at least one *Dictyostelium* disintegrin, AmpA, is needed throughout development for cell fate specification⁷⁰. We also identified distant relatives of vinculin and α -catenin—normally associated with adherens junctions—which support the idea that the epithelium-like sheet of cells that surrounds the stalk tube contains such junctions⁷¹. Consistent with this, the *Dictyostelium* genome encodes numerous proteins previously described as components of adherens junctions in Metazoa, such as β -catenin (Aardvark), α -actinin, formins, VASP and myosin VII.

In animals, tandem repeats of immunoglobulin, cadherin, fibronectin III or E-set domains are often present in cell adhesion proteins, although their common protein fold pre-dates the emergence of eukaryotes. EGF/laminin domains are also found in adhesion proteins but, before the analysis of the *Dictyostelium* genome, no non-metazoan was known to have more than two EGF repeats in a single predicted protein. *Dictyostelium* has 61 predicted proteins containing repeated E-set or EGF/laminin domains, and many of these contain additional domains that suggest they have roles in cell adhesion or cell recognition, such as mannose-6-phosphate receptor, fibronectin III, or growth factor receptor domains and transmembrane domains (Fig. 9). In support of this idea, four of these proteins (LagC, LagD, AmpA and ComC) have been shown to be required for cell adhesion and signalling during development^{70,72–74}.

Cellulose-based structures

During development, *Dictyostelium* cells produce a number of cellulose-based structural elements. *Dictyostelium* slugs synthesize an extracellular matrix, or sheath, around themselves that is comprised of proteins and cellulose. Several of the smaller sheath proteins bind cellulose and are believed to have a role in slug migration, whereas the larger, cysteine-rich EcmA protein is essential for full integrity of the sheath and for establishing correct slug shape^{75,76}. During terminal differentiation, cellulose is deposited in the stalk and in the cell walls of the stalk and spore cells^{77–79}. The first confirmed eukaryotic gene for cellulose synthase was discovered in *Dictyostelium* and this gene has since been recognized in many plants, *N. crassa* and the ascidian *Ciona intestinalis*⁸⁰. The fungal and urochordate enzymes are more closely related to the *Dictyostelium* homologue than to plant or bacterial cellulose synthases, indicating that the common ancestor of fungi and animals carried a gene for

cellulose synthase that was subsequently lost in most animals. The *Dictyostelium* genome encodes more than 40 additional proteins that are likely to be involved in cellulose synthesis or degradation, and are probably involved in the production and remodelling of cellulose fibres of the slug sheath, stalk tube and cell walls (see Supplementary Information).

The fundamental similarities in cellular cooperation found in *Dictyostelium* and in the Metazoa clearly resulted in a parallel positive selection for structural and regulatory genes required for cell motility, adhesion and signalling. *Dictyostelium* uses a set of signals and adhesion proteins that are distinct from those employed for similar purposes in Metazoa but, like the Metazoa, *Dictyostelium* has maintained a diversity of GPCRs, protein kinases and ABC transporters that enable it to respond to those signals. *Dictyostelium* has also retained and modified an organizational strategy perfected in plants, basing several structural elements on cellulose. At one

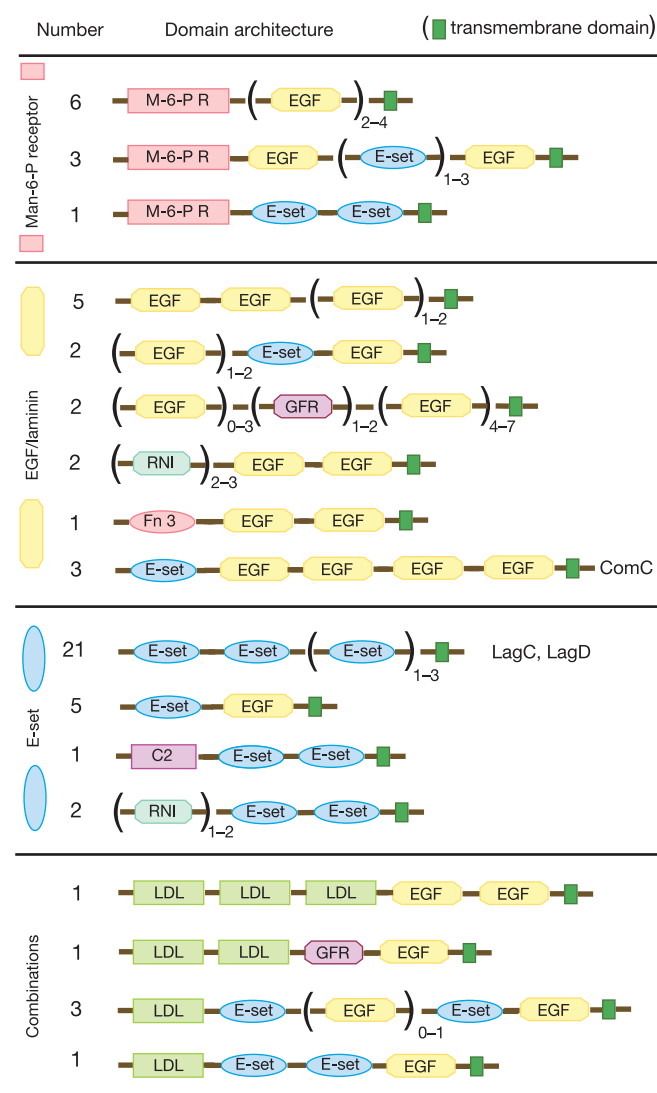


Figure 9 Putative adhesion/signalling proteins. Proteins containing repeated EGF/laminin and/or E-set SCOP Superfamily domains are classified into groups containing mannose-6-phosphate receptor, mainly EGF/laminin, mainly E-set, or combinations of domains. Most of these proteins have predicted transmembrane domains and so are expected to be cell surface proteins. ComC, LagC and LagD are proteins that have been characterized to have adhesion and/or signalling functions during multicellular development^{72–74}. C2, calcium-dependent lipid binding; Fn 3, fibronectin type III; GFR, growth factor receptor; LDL, L domain-like leucine-rich repeat; M-6-P R, mannose-6-phosphate receptor; RNI, RNI-like.

level *Dictyostelium* has achieved multicellularity by using strategies that are similar to plants and Metazoa, but the differences between them suggest convergent evolution, rather than lineal descent from an ancestor with overt or latent multicellular capacities.

Conclusion

The complete protein repertoire of *Dictyostelium* provides a new perspective for studying its cellular and developmental biology. At a systems level, *Dictyostelium* provides a level of complexity that is greater than the yeasts, but much simpler than plants or animals. Thus, high-resolution molecular analyses in this system may reveal control networks that are difficult to study in more complex systems, and may presage regulatory strategies used by higher organisms^{81–83}. At a practical level, the comparative genomics of *Dictyostelium* and related pathogens, such as *Entamoeba histolytica*, should aid in the functional definition of amoebozoan-specific genes that may open new avenues of research aimed at controlling amoebic diseases. *Dictyostelium*'s adeptness at hunting bacteria also renders it susceptible to infections by intracellular bacterial pathogens^{84,85}. *Dictyostelium* and human macrophages display fundamental similarities in their cell biology, which has spurred the use of *Dictyostelium* as a model host for bacterial pathogenesis. It is also an attractive model in which to study other disease processes: for a number of human disease-related proteins, it provides a test-bed for studying their functions in a model organism that has greater similarity to higher eukaryotes than do the yeasts, yet shares the latter's experimental tractability.

The high frequency of repeated amino acid tracts in *Dictyostelium* proteins has long been known anecdotally, but we can now survey their precise nature and number, and find them to be more abundant than in any other sequenced genome. Many human diseases result from the expansion of triplet nucleotide repeats, some of which encode polyglutamine tracts that cause cell degeneration^{86,87}. Learning how *Dictyostelium* cells tolerate so many proteins with amino acid homopolymers will, we hope, help to elucidate the roles of these motifs in protein function and dysfunction.

Comparative genomic studies in eukaryotes are providing the raw material for global examinations of the evolution of cellular regulation and developmental mechanisms⁸⁸. Many genes have been lost in one species but retained in others, such that each new genome sequence adds to our understanding of the genetic complement of the eukaryotic progenitor. Thus, our understanding of eukaryotes will continue to be refined as more genome sequences become available from representatives of large groups of organisms whose genomes remain largely unexplored, such as the amoebozoan. The surprising molecular diversity of the *Dictyostelium* proteome, which includes protein assemblages usually associated with fungi, plants or animals, suggests that their last common ancestor had a greater number of genes than had been previously appreciated. □

Methods

Details on the availability of reagents can be found in the Supplementary Information. All analyses described here were performed on Version 2.0 of the genome sequence. Updates to the sequence and annotation are available at <http://www.dictybase.org> and <http://www.genedb.org/genedb/dicty/index.jsp>. Further details of analyses not explicitly described below can be found in the Supplementary Information.

HAPPY mapping

A short-range (~100-kb), high-resolution (±8.54-kb) mapping panel was prepared as described⁹. Briefly, 96 aliquots each containing ±0.52 haploid genome equivalents of sheared AX4 genomic DNA were pre-amplified by PEP (primer extension pre-amplification⁸⁹). A total of 4,913 STS markers (Supplementary Table 1) were typed by two-phase hemi-nested polymerase chain reaction (PCR; multiplexed for up to 1,200 markers in the first phase) on aliquots of the diluted PEP products. Maps were assembled from good-quality data essentially as described previously⁹. A second, longer-range (±150 kb) mapping panel was used to confirm some linkages on chromosomes 2 and 5. HAPPY map analysis and PCR primer design for HAPPY mapping was performed using various custom programs (P.H.D. and A.T.B., unpublished).

Chromosome purification

Genomic DNA from *D. discoideum* strain AX4 was prepared and separated by pulsed field gel electrophoresis essentially as described^{27,9}, except that gels were run in stacked pairs; one member of each pair was stained with ethidium bromide, and bands excised from its unstained counterpart by ligation.

WCS and YAC subclone libraries

For WCS libraries, gel slices (above) were disrupted by several passages through a 30-gauge syringe needle, digested with β-agarase (NEB) and phenol-extracted. DNA was concentrated by ethanol precipitation, sonicated, end-blunted using mung bean nuclease and size-fractionated on 0.8% low-melting-point agarose gels. Fractions of 1.4–2 kb and 2–4 kb were excised, DNA extracted as before and ligated into the *Sma*I site of pUC18 or pUC19. Clone propagation and template preparation followed standard protocols.

For YAC subclone libraries, AX4-derived YACs were identified (and their position and integrity confirmed) by screening the set described by ref. 22 using markers from the HAPPY map. Subclones were prepared from PFG-purified YACs essentially as for the WCS libraries; contaminating yeast-derived sequences were filtered out *in silico*.

Sequencing and assembly

Details of the sequencing and assembly methods can be found in Supplementary Information. Generally, mapped sequence features were used to nucleate sequence contigs assembled from the WCS data, and extended using read-pair information and iterative searches for overlapping sequences, followed by directed gap closure using a range of approaches.

Fluorescent *in situ* hybridization

In situ hybridization was performed as in ref. 17.

Gene prediction and identification of sequence features

Full details are provided in the Supplementary Information. Briefly, automated gene prediction was performed using a combination of programs that had been trained on well-characterized *D. discoideum* genes, and the results integrated with reference to *D. discoideum* complementary DNA sequences and homology to genes in other species. Other features in the predicted proteins, and other sequence features, were identified using a variety of software packages.

Analysis of functional gene clustering

Microarray targets (refs 53, 90, 91; and N. Van Driessche and G. Shaulsky, unpublished data) and gene models were mapped onto the genome sequence using BLAST⁹² and the modified LIS algorithm⁹³. To look for clustering of genes with correlated temporal expression profiles, pairwise correlation coefficients were calculated for genes with known expression profiles on each chromosome⁹¹. Blocks of ≥6 consecutive genes were sought, for which either (1) all pairwise correlation coefficients were positive and ≥70% were >0.2 (genes with similar developmental trajectories) or (2) each gene had a partner with an absolute correlation coefficient value of >0.6 (tightly co-regulated genes); no statistically significant clusters met these criteria.

To look for clustering of genes associated with specific developmental stages^{94,95} or cell types^{90,96}, the genome was scanned with various sized windows⁹⁷ for regions with significant ($P < 0.01$) over-representation of genes in any one of these groups.

Analysis of duplicated genes

Predicted protein sequences were clustered using TribeMCL⁹⁸, using a BLASTP expectation of $<10^{-40}$ as a cutoff. A χ^2 test invalidated the hypothesis that members of a family are randomly distributed in the genome. Within each family, protein divergences (similarity distances computed using the 'ProtDist' module of PHYLIP; <http://evolution.genetics.washington.edu/phylip.html>) and physical intergenic distances between all pairs of family members were tabulated, and the correlation coefficient between the former and latter values was calculated. Analysis was performed on the 86 gene families (representing 155 gene pairs) with at least 10 intrachromosomal distance pairings to provide robust statistical confidence.

Other sequence analyses and graphical representation

Other sequence analyses (nucleotide and dinucleotide composition; identification of simple-sequence repeats in nucleotide and protein sequence; coding density computation; tRNA cluster identification) were performed using a range of custom software (P.H.D. and A.T.B., unpublished). Graphical representation of chromosomes in Fig. 2 was done primarily using Cinema4D-8.5 (Maxon Computer GmbH) after pre-processing using custom software (P.H.D.).

Received 16 September 2004; accepted 17 February 2005; doi:10.1038/nature03481.

1. Kessin, R. H. *Dictyostelium—Evolution, Cell Biology, and the Development of Multicellularity*, xiv, 294 (Cambridge Univ. Press, Cambridge, 2001).
2. Konijn, T. M. *et al.* The acrasin activity of adenosine-3',5'-cyclic phosphate. *Proc. Natl Acad. Sci. USA* **58**, 1152–1154 (1967).
3. Müller, K. & Gerisch, G. A specific glycoprotein as the target site of adhesion blocking Fab in aggregating *Dictyostelium* cells. *Nature* **274**, 445–449 (1978).
4. Raper, K. B. Pseudopodium formation and organization in *Dictyostelium discoideum*. *J. Elisha Mitchell Sci. Soc.* **56**, 241–282 (1940).

5. Raper, K. B. *Dictyostelium discoideum*, a new species of slime mold from decaying forest leaves. *J. Agr. Res.* **50**, 135–147 (1935).
6. Knecht, D. A., Cohen, S. M., Loomis, W. F. & Lodish, H. F. Developmental regulation of *Dictyostelium discoideum* actin gene fusions carried on low-copy and high-copy transformation vectors. *Mol. Cell. Biol.* **6**, 3973–3983 (1986).
7. Dear, P. H. & Cook, P. R. HAPPY mapping—linkage mapping using a physical analog of meiosis. *Nucleic Acids Res.* **21**, 13–20 (1993).
8. Konfortov, B. A., Cohen, H. M., Bankier, A. T. & Dear, P. H. A high-resolution HAPPY map of *Dictyostelium discoideum* chromosome 6. *Genome Res.* **10**, 1737–1742 (2000).
9. Glöckner, G. *et al.* Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. *Nature* **418**, 79–85 (2002).
10. Urushihara, H. *et al.* Analyses of cDNAs from growth and slug stages of *Dictyostelium discoideum*. *Nucleic Acids Res.* **32**, 1647–1653 (2004).
11. Smith, S. S. & Ratner, D. I. Lack of 5-methylcytosine in *Dictyostelium discoideum* DNA. *Biochem. J.* **277**, 273–275 (1991).
12. Glöckner, G. *et al.* The complex repeats of *Dictyostelium discoideum*. *Genome Res.* **11**, 585–594 (2001).
13. Crick, F. H. Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* **19**, 548–555 (1966).
14. Soll, D. & Rajbhandary, U. (ed.) *tRNA: Structure, Biosynthesis and Function* (ASM, Washington DC, 1995).
15. Burger, G., Plante, I., Lonergan, K. M. & Gray, M. W. The mitochondrial DNA of the amoeboid protozoan, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. *J. Mol. Biol.* **3**, 522–537 (1995).
16. Ogawa, S. *et al.* The mitochondrial DNA of *Dictyostelium discoideum*: complete sequence, gene content and genome organization. *Mol. Gen. Genet.* **263**, 514–519 (2000).
17. Sugcan, R. *et al.* Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in *Dictyostelium*. *Nucleic Acids Res.* **31**, 2361–2368 (2003).
18. Szafranski, K., Dingermann, T., Glöckner, G. & Winckler, T. Template jumping by a LINE reverse transcriptase has created a SINE-like 5S rRNA retropseudogene in *Dictyostelium*. *Mol. Genet. Genom.* **271**, 98–102 (2004).
19. Eichler, E. E. & Sankoff, D. Structural dynamics of eukaryotic chromosome evolution. *Science* **301**, 793–797 (2003).
20. Cappello, J., Cohen, S. M. & Lodish, H. F. *Dictyostelium* transposable element DIRS-1 preferentially inserts into DIRS-1 sequences. *Mol. Cell. Biol.* **4**, 2207–2213 (1984).
21. Cappello, J., Handelsman, K. & Lodish, H. F. Sequence of *Dictyostelium* DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. *Cell* **43**, 105–115 (1985).
22. Loomis, W. F., Welker, D., Hughes, J., Maghakian, D. & Kuspa, A. Integrated maps of the chromosomes in *Dictyostelium discoideum*. *Genetics* **141**, 147–157 (1995).
23. Goodwin, T. J. & Poulter, R. T. Multiple LTR-retrotransposon families in the asexual yeast *Candida albicans*. *Genome Res.* **10**, 174–191 (2000).
24. Appelgren, H., Kniola, B. & Ekwall, K. Distinct centromere domain structures with separate functions demonstrated in live fission yeast cells. *J. Cell Sci.* **116**, 4035–4042 (2003).
25. Kuspa, A., Maghakian, D., Bergesch, P. & Loomis, W. F. Physical mapping of genes to specific chromosomes in *Dictyostelium discoideum*. *Genomics* **13**, 49–61 (1992).
26. McClintock, B. The production of homozygous deficient tissues with mutant characteristics by means of the aberrant mitotic behaviour of ring-shaped chromosomes. *Genetics* **23**, 315–376 (1938).
27. Cox, E. C., Vocke, C. D., Walter, S., Gregg, K. Y. & Bain, E. S. Electrophoretic karyotype for *Dictyostelium discoideum*. *Proc. Natl Acad. Sci. USA* **87**, 8247–8251 (1990).
28. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–941 (2001).
29. Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
30. Trusov, Y. A. & Dear, P. H. A molecular clock based on the expansion of gene families. *Nucleic Acids Res.* **24**, 995–999 (1996).
31. Kellis, M., Birren, B. W. & Lander, E. S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617–624 (2004).
32. Dujon, B. *et al.* Genome evolution in yeasts. *Nature* **430**, 35–44 (2004).
33. Raper, K. B. *The Dictyostelids* (Princeton Univ. Press, Princeton, New Jersey, 1984).
34. Loomis, W. F. & Smith, D. W. Consensus phylogeny of *Dictyostelium*. *Experientia* **51**, 1110–1115 (1995).
35. Baldauf, S. L. & Doolittle, W. F. Origin and evolution of the slime molds (Mycetozoa). *Proc. Natl Acad. Sci. USA* **94**, 12007–12012 (1997).
36. Baptiste, E. *et al.* The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc. Natl Acad. Sci. USA* **99**, 1414–1419 (2002).
37. Nara, T., Hshimoto, T. & Aoki, T. Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. *Gene* **257**, 209–222 (2000).
38. Olsen, R. & Loomis, W. F. A model of orthologous protein sequence divergence. *J. Mol. Evol.* (in the press).
39. Thomason, P. & Kay, R. Eukaryotic signal transduction via histidine-aspartate phosphorelay. *J. Cell Sci.* **113**, 3141–3150 (2000).
40. Fortini, M. E., Skupski, M. P., Boguski, M. S. & Hariharan, I. K. A survey of human disease gene counterparts in the *Drosophila* genome. *J. Cell Biol.* **150**, F23–F30 (2000).
41. Jain, R., Rivera, M. C., Moore, J. E. & Lake, J. A. Horizontal gene transfer in microbial genome evolution. *Theor. Popul. Biol.* **61**, 489–495 (2002).
42. Richards, T. A., Hirt, R. P., Williams, B. A. & Embley, T. M. Horizontal gene transfer and the evolution of parasitic protozoa. *Protist* **154**, 17–32 (2003).
43. Iyer, L. M., Aravind, L., Coon, S. L., Klein, D. C. & Koonin, E. V. Evolution of cell-cell signaling in animals: did late horizontal gene transfer from bacteria have a role? *Trends Genet.* **20**, 292–299 (2004).
44. Myllykallio, H. *et al.* An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297**, 105–107 (2002).
45. Kessin, R. H., Gundersen, G. G., Zaydfudim, V., Grimson, M. & Blanton, R. L. How cellular slime molds evade nematodes. *Proc. Natl Acad. Sci. USA* **93**, 4857–4861 (1996).
46. Morris, H. R., Taylor, G. W., Masento, M. S., Jermyn, K. A. & Kay, R. R. Chemical structure of the morphogen differentiation inducing factor from *Dictyostelium discoideum*. *Nature* **328**, 811–814 (1987).
47. Cane, D. E., Walsh, C. T. & Khosla, C. Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63–68 (1998).
48. Holland, I. B. & Blight, M. A. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J. Mol. Biol.* **293**, 381–399 (1999).
49. Andrade, A. C., Van Nistelrooy, J. G., Peery, R. B., Skatrud, P. L. & De Waard, M. A. The role of ABC transporters from *Aspergillus nidulans* in protection against cytotoxic agents and in antibiotic production. *Mol. Gen. Genet.* **263**, 966–977 (2000).
50. Mendez, C. & Salas, J. A. The role of ABC transporters in antibiotic-producing organisms: drug secretion and resistance mechanisms. *Res. Microbiol.* **152**, 341–350 (2001).
51. Schoonbeek, H. J., Raaijmakers, J. M. & De Waard, M. A. Fungal ABC transporters and microbial interactions in natural environments. *Mol. Plant Microbe Interact.* **15**, 1165–1172 (2002).
52. Anjard, C. & Loomis, W. F. Evolutionary analyses of ABC transporters of *Dictyostelium discoideum*. *Eukaryot. Cell* **1**, 643–652 (2002).
53. Iranfar, N., Fuller, D. & Loomis, W. F. Genome-wide expression analyses of gene regulation during early development of *Dictyostelium discoideum*. *Eukaryot. Cell* **2**, 664–670 (2003).
54. Devreotes, P. N. & Zigmond, S. H. Chemotaxis in eukaryotic cells: A focus on leukocytes and *Dictyostelium*. *Annu. Rev. Cell Biol.* **4**, 649–686 (1988).
55. Noegel, A. A. & Schleicher, M. The actin cytoskeleton of *Dictyostelium*: a story told by mutants. *J. Cell Sci.* **113**, 759–766 (2000).
56. Rivero, F. & Somesh, B. P. Signal transduction pathways regulated by Rho GTPases in *Dictyostelium*. *J. Muscle Res. Cell Motil.* **23**, 737–749 (2002).
57. Merlot, S. & Firtel, R. A. Leading the way: directional sensing through phosphatidylinositol 3-kinase and other signalling pathways. *J. Cell Sci.* **116**, 3471–3478 (2003).
58. Bockaert, J. & Pin, J. P. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J.* **18**, 1723–1729 (1999).
59. Ginsburg, G. T. *et al.* The regulation of *Dictyostelium* development by transmembrane signalling. *J. Eukaryot. Microbiol.* **42**, 200–205 (1995).
60. Raisley, B., Zhang, M., Hereld, D. & Hadwiger, J. A. A cAMP receptor-like G protein-coupled receptor with roles in growth regulation and development. *Dev. Biol.* **265**, 433–445 (2004).
61. King, N., Hittinger, C. T. & Carroll, S. B. Evolution of key cell signaling and adhesion protein families predates animal origins. *Science* **301**, 361–363 (2003).
62. Kawata, T. *et al.* SH2 signaling in a lower eukaryote: A STAT protein that regulates stalk cell differentiation in *Dictyostelium*. *Cell* **89**, 909–916 (1997).
63. Thien, C. B. & Langdon, W. Y. Cbl: many adaptations to regulate protein tyrosine kinases. *Nature Rev. Mol. Cell Biol.* **2**, 294–307 (2001).
64. Meng, W. *et al.* Structure of the amino-terminal domain of Cbl complexed to its binding site on ZAP-70 kinase. *Nature* **398**, 84–90 (1999).
65. Shaulsky, G., Kuspa, A. & Loomis, W. F. A multidrug resistance transporter serine protease gene is required for prestalk specialization in *Dictyostelium*. *Genes Dev.* **9**, 1111–1122 (1995).
66. Anjard, C., Zeng, C., Loomis, W. F. & Nellen, W. Signal transduction pathways leading to spore differentiation in *Dictyostelium discoideum*. *Dev. Biol.* **193**, 146–155 (1998).
67. Good, J. R. *et al.* TagA, a putative serine protease/ABC transporter of *Dictyostelium* that is required for cell fate determination at the onset of development. *Development* **130**, 2953–2965 (2003).
68. Good, J. R. & Kuspa, A. Evidence that a cell-type-specific efflux pump regulates cell differentiation in *Dictyostelium*. *Dev. Biol.* **220**, 53–61 (2000).
69. Chibalina, M. V., Anjard, C. & Insall, R. H. Gdt2 regulates the transition of *Dictyostelium* cells from growth to differentiation. *BMC Dev. Biol.* **4**, 8 (2004).
70. Blumberg, D. D., Ho, H. N., Petty, C. L., Varney, T. R. & Gandham, S. AmpA, a modular protein containing disintegrin and ornatin domains, has multiple effects on cell adhesion and cell fate specification. *J. Muscle Res. Cell Motil.* **23**, 817–828 (2002).
71. Grimson, M. J. *et al.* Adherens junctions and β -catenin-mediated cell signalling in a non-metazoan organism. *Nature* **408**, 727–731 (2000).
72. Dynes, J. L. *et al.* LagC is required for cell-cell interactions that are essential for cell-type differentiation in *Dictyostelium*. *Genes Dev.* **8**, 948–958 (1994).
73. Wang, J. *et al.* The membrane glycoprotein gp150 is encoded by the *lagC* gene and mediates cell-cell adhesion by heterophilic binding during *Dictyostelium* development. *Dev. Biol.* **227**, 734–745 (2000).
74. Kibler, K., Svetz, J., Nguyen, T. L., Shaw, C. & Shaulsky, G. A cell-adhesion pathway regulates intercellular communication during *Dictyostelium* development. *Dev. Biol.* **264**, 506–521 (2003).
75. Morrison, A. *et al.* Disruption of the gene encoding the EcmA, extracellular matrix protein of *Dictyostelium* alters slug morphology. *Dev. Biol.* **163**, 457–466 (1994).
76. Wang, Y. Z., Slade, M. B., Gooley, A. A., Atwell, B. J. & Williams, K. L. Cellulose-binding modules from extracellular matrix proteins of *Dictyostelium discoideum* stalk and sheath. *Eur. J. Biochem.* **268**, 4334–4345 (2001).
77. Freeze, H. & Loomis, W. F. Chemical analysis of stalk components of *Dictyostelium discoideum*. *Biochim. Biophys. Acta* **539**, 529–537 (1978).
78. Zhang, P., McGlynn, A., Loomis, W. F., Blanton, R. L. & West, C. M. Spore coat formation and timely sporulation depend on cellulose in *Dictyostelium*. *Differentiation* **67**, 72–79 (2001).
79. West, C. M., Zhang, P., McGlynn, A. C. & Kaplan, L. Outside-in signaling of cellulose synthesis by a spore coat protein in *Dictyostelium*. *Eukaryot. Cell* **1**, 281–292 (2002).
80. Blanton, R. L., Fuller, D., Iranfar, N., Grimson, M. J. & Loomis, W. F. The cellulose synthase gene of *Dictyostelium*. *Proc. Natl Acad. Sci. USA* **97**, 2391–2396 (2000).
81. Thomason, P. A. *et al.* An intersection of the cAMP/PKA and two-component signal transduction systems in *Dictyostelium*. *EMBO J.* **17**, 2838–2845 (1998).
82. Maeda, M. *et al.* Periodic signaling controlled by an oscillatory circuit that includes protein kinases ERK2 and PKA. *Science* **304**, 875–878 (2004).
83. Soler-Lopez, M. *et al.* Structure of an activated *Dictyostelium* STAT in its DNA-unbound form. *Mol. Cell* **13**, 791–804 (2004).
84. Solomon, J. M., Rupper, A., Cardelli, J. A. & Isberg, R. R. Intracellular growth of *Legionella pneumophila* in *Dictyostelium discoideum*, a system for genetic analysis of host-pathogen interactions. *Infect. Immun.* **68**, 2939–2947 (2000).
85. Skriwan, C. *et al.* Various bacterial pathogens and symbionts infect the amoeba *Dictyostelium discoideum*. *Int. J. Med. Microbiol.* **291**, 615–624 (2002).
86. Zoghbi, H. Y. & Orr, H. T. Glutamine repeats and neurodegeneration. *Annu. Rev. Neurosci.* **23**, 217–247 (2000).

87. Brown, L. Y. & Brown, S. A. Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends Genet.* **20**, 51–58 (2004).
88. Rubin, G. M. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215 (2000).
89. Zhang, L. *et al.* Whole genome amplification from a single cell: Implications for genetic analysis. *Proc. Natl Acad. Sci. USA* **89**, 5847–5851 (1992).
90. Iranfar, N. *et al.* Expression patterns of cell-type-specific genes in *Dictyostelium*. *Mol. Biol. Cell* **12**, 2590–2600 (2001).
91. Van Driessche, N. *et al.* A transcriptional profile of multicellular development in *Dictyostelium discoideum*. *Development* **129**, 1543–1552 (2002).
92. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
93. Zhang, H. Alignment of BLAST high-scoring segment pairs based on the longest increasing subsequence algorithm. *Bioinformatics* **19**, 1391–1396 (2003).
94. Katoh, M. *et al.* An orderly retreat: dedifferentiation is a regulated process. *Proc. Natl Acad. Sci. USA* **101**, 7005–7010 (2004).
95. Xu, Q. *et al.* Transcriptional transitions during *Dictyostelium* spore germination. *Eukaryot. Cell* **3**, 1101–1110 (2004).
96. Maeda, M. *et al.* Changing patterns of gene expression in *Dictyostelium* prestalk cell subtypes recognized by *in situ* hybridization with genes from microarray analyses. *Eukaryot. Cell* **2**, 627–637 (2003).
97. Cohen, B. A., Mitra, R. D., Hughes, J. D. & Church, G. M. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nature Genet.* **26**, 183–186 (2000).
98. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements Sequencing and analysis of chromosomes 1, 2 and 3 were supported by grants from the DFG and by Köln Fortune, and that of chromosomes 4, 5 and 6 in the USA by grants from NICHD/NIH. Work in the UK/European Union (EU) was supported by a

programme grant from the MRC to J.W., R.R.K., B.B. and P.H.D., and by the EU. Analyses at dictyBase were supported by a grant from the NIGMS/NIH to R.L.C. The *Dictyostelium* cDNA project was supported by Research for the Future of JSPS and by Grants-in-Aid for Scientific Research on Priority Areas of MEXT of Japan. The German team wishes to thank S. Förste, N. Zeisse, S. Rothe, S. Landmann, R. Schultz, C. Neuhoff and R. Müller for technical assistance. The US team thanks S. Kaminsky, S. Klein and T. Hewitt for their scientific foresight in the early stages of this project and for their support of *Dictyostelium* as a model system, and H. Hosak, O. Delgado, L. Lewis, K. Hamilton, J. Hume, C. Kovar Smith, D. Neal, P. Havlak, K. J. Durbin and P. Burch of the HGSC at Baylor College of Medicine. R.S. thanks L. Cortez, E. Joyner and B. Hill for their assistance during the course of the project. C.B. thanks M. Veron and P. Glaser for their support and discussions. P.H.D. thanks H. O'Hare for early involvement in the project, A. Ivens for discussions, and the MRC Centre Visual Aids department, Cambridge, for advice on graphics. We also thank S. Bowman and D. Lawson for their contribution to the EUDICT region in the initial stages of the project, and D. Martin at the Wellcome Trust Biocentre, University of Dundee, for running the GOTcha search of our gene models. The Japanese cDNA project thanks N. Ogasawara and I. Takeuchi for comments and encouragement, and others who participated in earlier stages of the project.

Author contributions M. Platzer, R. R. Kay, J. Williams, P. H. Dear, A. A. Noegel, B. Barrell and A. Kuspa are co-senior authors.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.H.D. (phd@mrc-lmb.cam.ac.uk). Sequence data for the genome were deposited in the GenBank nucleotide database under the project accession number AAFI00000000. The six chromosomal assemblies have accession numbers CM000150–CM000155, and their component sequence contigs have accession numbers AAFI01000001–AAFI01000336.

Lipoprotein particles are required for Hedgehog and Wingless signalling

Daniela Panáková*, Hein Sprong*†, Eric Marois, Christoph Thiele & Suzanne Eaton

Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse-108, 01307 Dresden, Germany

* These authors contributed equally to this work

† Present address: Department of Membrane Enzymology, Institute of Biomembranes, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Wnt and Hedgehog family proteins are secreted signalling molecules (morphogens) that act at both long and short range to control growth and patterning during development. Both proteins are covalently modified by lipid, and the mechanism by which such hydrophobic molecules might spread over long distances is unknown. Here we show that Wingless, Hedgehog and glycosphosphatidylinositol-linked proteins copurify with lipoprotein particles, and co-localize with them in the developing wing epithelium of *Drosophila*. In larvae with reduced lipoprotein levels, Hedgehog accumulates near its site of production, and fails to signal over its normal range. Similarly, the range of Wingless signalling is narrowed. We propose a novel function for lipoprotein particles, in which they act as vehicles for the movement of lipid-linked morphogens and glycosphosphatidylinositol-linked proteins.

In the developing wing of *Drosophila*, Hedgehog activates short-range target gene expression up to five cells away from its source of production, and longer-range targets over more than twelve cell diameters¹. Wingless can signal through a range of over 30 cell diameters². These morphogens are anchored to the membrane via covalent lipid modification^{3–9}. The mechanisms that allow long-range movement of molecules with such strong membrane affinity are unclear.

Like Wingless and Hedgehog, glycosphosphatidylinositol (gpi)-linked proteins transfer between cells with their lipid anchor intact^{10–12}. We observed that gpi-linked green fluorescent protein (GFP) expressed in Wingless-producing cells spreads into receiving tissue at the same rate as Wingless, where it co-localizes with Wingless in endosomes. Thus, we proposed that these proteins travel together on a membranous particle, which we called an argosome¹³. How might argosomes form? One possibility is that argosomes are membranous exovesicles. Such particles could be generated by plasma membrane vesiculation, or by an exosome-related mechanism¹⁴. Alternatively, argosomes might resemble lipoprotein particles like low-density lipoprotein (LDL). Vertebrate lipoprotein particles are scaffolded by apolipoproteins and comprise a phospholipid monolayer surrounding a core of esterified cholesterol and triglyceride. Insects construct similar particles called lipophorins^{15,16}. Lipid-modified proteins of the exoplasmic face of the membrane (such as GFPgpi, Wingless or Hedgehog) might insert into the outer phospholipid monolayer of such a particle via their attached lipid moieties. Here, we use biochemical fractionation to determine the sort of particle with which lipid-linked proteins associate, and genetic means to address its function.

Lipid-linked proteins copurify with lipophorin

We compared sedimentation of Wingless, Hedgehog and gpi-linked proteins to that of transmembrane proteins, exosomes and lipophorin particles. To mark exosomes, we used flies expressing a vertebrate CD63:GFP fusion construct. CD63 is a tetraspanin that localizes to internal vesicles of multivesicular endosomes, and is released on exosomes^{17,18}. In *Drosophila* imaginal discs, CD63:GFP localizes to late endosomes in producing cells, consistent with vertebrate studies (Supplementary Fig. S1A–D). It is released and endocytosed by neighbouring cells between one and three cell diameters away (Supplementary Fig. S1A,E), indicating that it is present on exosomes.

To mark lipoprotein particles, we made antibodies to *Drosophila*

apolipophorins I and II (ApoLI and ApoLII); these proteins are generated by cleavage of the precursor pro-Apolipophorin^{19,20}. Lipophorin is produced in the fat body²⁰; consistent with this, we cannot detect *apolipophorin* transcripts in imaginal discs (data not shown). Nevertheless, the ApoLI and ApoLII proteins are as abundant in discs as in the fat body (Supplementary Fig. S1F and G).

Plasma membrane and exosomal markers are completely pelleted after centrifugation for 3 h at 120,000g, whereas most ApoLII remains in the supernatant (Fig. 1a). Most Wingless:GFP and Hedgehog is present in the pellet, as are the gpi-linked proteins Fasciclin I²¹, Connectin²², Klingon²³ and Acetylcholinesterase²⁴ (Fig. 1b); this is not unexpected, because these proteins localize to the plasma membrane and internal membrane compartments. Surprisingly, however, some Wingless:GFP (6%), Hedgehog (2%) and gpi-linked proteins (14–22%) remain in the supernatant (Fig. 1b).

The 120,000g supernatant (S120) contains both free soluble proteins and lipoprotein particles. To separate them, we performed isopycnic density centrifugation. In these gradients, lipophorin moves to the top low-density fraction whereas soluble proteins are present in higher-density fractions (top two panels of Fig. 1c). Gpi-linked proteins are found almost entirely in the top fraction with lipophorin. Treating the S120 with Phosphatidylinositol-specific phospholipase C (PI-PLC) before density centrifugation shifts their migration to higher-density fractions (Fig. 1c). This suggests that gpi-linked proteins associate with low-density particles via their gpi anchor.

Similarly, when S120s from larvae that express Wingless:GFP or Hedgehog:HA in imaginal discs are subjected to isopycnic density centrifugation, these proteins are found in the lowest-density fraction with ApoLII, as is endogenous Hedgehog (Fig. 1d). Antibodies to endogenous Wingless detect a doublet in the top fraction and a band of somewhat higher mobility in high-density fractions. These data indicate that non-membrane-bound Wingless and Hedgehog associate with low-density particles in imaginal discs *in vivo*; other larval tissues may secrete Wingless in a non-lipophorin-associated form.

We worried that lipophorin particles in the haemolymph might extract proteins from discs during larval homogenization, so we repeated these experiments using dissected discs. All Wingless, Hedgehog and ApoLII in the imaginal disc S120s are present on low-density particles (Fig. 1e), suggesting that their association is not an artefact of homogenization. Consistent with this, incubating

pelleted imaginal disc membranes with an excess of purified lipoprotein particles does not extract Hedgehog:HA from membranes under the conditions used for homogenization (Fig. 1f). This suggests that association of lipid-linked morphogens with lipophorin depends on active cellular processes and does not occur during extract preparation.

To ask whether lipid-linked proteins associated with lipophorin, or with some other low-density particle, we immunoprecipitated ApoLII from larval S120s and probed precipitates for Wingless, Hedgehog or GFPgpi. These proteins are immunoprecipitated by anti-ApoLII, but not pre-immune serum (Fig. 1g). Furthermore, anti-ApoLII is unable to precipitate secreted GFP that does not contain a gpi anchor (Fig. 1h). Hedgehog and Fas-1 also immunoprecipitated with ApoLII from the more purified top fraction of KBr gradients (Supplementary Fig. S1). Thus, lipid-linked morphogens and gpi-linked proteins associate directly with lipophorin particles.

Morphogens colocalize with lipophorin

These experiments do not exclude the possibility that some Wingless or Hedgehog in the 120,000g pellet (P120) might be present on exosomes. To investigate this, we expressed CD63:GFP in either Wingless- or Hedgehog- producing cells and looked for colocalization with CD63:GFP-labelled exosomes in receiving tissue. No significant co-localization is detected (Fig. 2d–f, j–l). Thus, it seems unlikely that imaginal disc cells release Wingless or Hedgehog on exosomes, although the mechanism remains a possibility for transmembrane ligands such as Boss or Notch^{25,26}.

To test whether Wingless or Hedgehog co-localized with lipoprotein particles, we incubated imaginal discs with purified lipophorin particles fluorescently labelled with Alexa488; although they work for western blotting, neither anti-ApoLI nor ApoLII antibodies detect endogenous lipophorin by immunofluorescence. Immunostaining reveals that Wingless and Hedgehog are found in the same endosomes as Alexa488lipophorin (Fig. 2a–c, g–i). Unsurprisingly, lipophorin uptake is not limited to areas where

these morphogens are abundant; lipophorin has a nutritional function as well and many potential receptors are encoded in the genome²⁷. Strong co-localization between lipophorin and lipid-linked morphogens is predicted if Wingless and Hedgehog are endocytosed with lipophorin. Nevertheless, we cannot exclude the possibility that these proteins were internalized separately and converged in the same endosomes.

Lipophorin–RNAi perturbs lipid transport

To assess the role of lipophorin in larval growth and development, we reduced the levels of ApoLI and II by RNA interference directed against two different regions of the *apolipophorin* messenger RNA. Similar phenotypes were produced by each construct. To express double-stranded (ds)RNA, we used a modified GAL4:UAS system in which expression of inverted repeats can be temporally controlled by heat-shock-dependent excision of an intervening HcRed cassette by the flippase (FLP) recombinase. We tested extracts from wild-type larvae or larvae harbouring *hs-flp*, GAL4 driver and UAS dsRNA constructs at various times after heat shock to see how fast lipophorin levels were reduced (Fig. 3a). Larvae of the latter genotype made only 50% of the wild-type level of ApoLII, even in the absence of heat shock; basal activity of the heat-shock promoter in the fat body causes HcRed excision in approximately 50% of fat-body cells, although excision strictly depends on heat shock in other larval tissues (data not shown). Although they survive less frequently, these flies have no obvious phenotype.

After heat shock, all fat-body cells excise the HcRed cassette and ApoLII levels decrease further. After four days, ApoLII is reduced to 5% of wild-type levels. ApoLI levels are reduced with similar kinetics (Supplementary Fig. S3). These animals prolong the third larval instar and rarely pupariate. We performed all the experiments described below on third-instar larvae 4–6 days after heat shock.

To investigate the requirement for lipophorin in lipid transport, we assessed the accumulation of neutral lipids in larval tissues by staining them with Nile Red²⁸. Cells of the posterior midgut

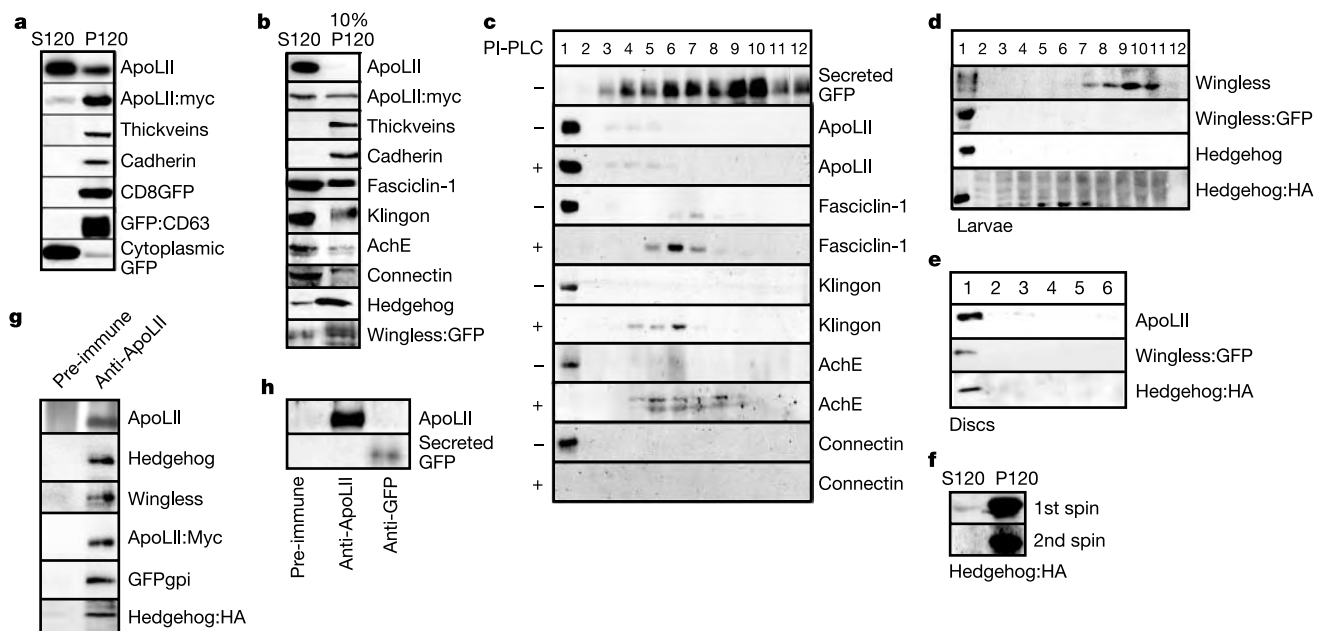


Figure 1 Lipid-linked proteins co-fractionate with lipophorin. Western blots of fractionated extracts probed with antibodies to indicated proteins. **a, b**, Larval S120s and indicated proportions of larval P120s. AchE, acetylcholinesterase. **c–e**, KBr isopycnic density gradient fractions made from larval (**c, d**) or disc (**e**) S120s. Top fraction, 1.14 g cm⁻³; bottom fraction, 1.4 g cm⁻³. +, PI-PLC-treated, –, mock-treated. **f**, Top

panel: P120 and S120 from Hedgehog:haemagglutinin (HA)-expressing discs, probed with anti-HA. Lower panel: P120 and S120 after incubating P120 at 4 °C with fivefold excess of purified lipoprotein particles and recentrifuging. **g, h**, S120s from wild-type or fusion-protein-expressing larvae immunoprecipitated with pre-immune, anti-ApoLII or anti-GFP serum.

normally contain many small lipid droplets (Fig. 3b). Lipophorin reduction causes a dramatic expansion of these droplets (Fig. 3c), suggesting that lipophorin is required for the efficient extraction of lipid from the midgut.

The wild-type fat body contains both small and large lipid droplets (Fig. 3d). Fat bodies of lipophorin-RNAi larvae are reduced in size and have fewer small lipid droplets (Fig. 3e), although larger droplets appeared normal. These data suggest that lipophorin delivers lipid to the fat body.

Lipid droplets in discs from lipophorin-RNAi larvae are fewer and smaller than in the wild type (compare Fig. 3f and g). Their discs are also reduced in size, particularly in the wing pouch (data not shown). Thus, discs require lipophorin for accumulation of lipid droplets and for growth. Neither Caspase3 activation nor membrane phosphatidylinositol 3,4,5-phosphate (PIP₃) accumulation is altered in lipophorin-RNAi discs (Supplementary Fig. S4), suggesting that their small size is not due to cell death or reduced insulin signalling²⁹.

Hedgehog function requires lipophorin

To test whether lipophorin association was required for Hedgehog

function, we examined Hedgehog distribution and signalling in lipophorin-RNAi larval discs. In wild-type discs, Hedgehog expressed in the posterior compartment moves across the anterior-posterior (AP) compartment boundary and activates transcription of short and long-range target genes. Cells closest to the source respond by activating the transcription of *collier* (Fig. 4b, c) and *patched* (Fig. 4h, i). Further away, Hedgehog activates transcription of *decapentaplegic* (Fig. 4a, b)^{1,30,31}. We monitored levels of Collier and a *decapentaplegic* reporter construct (*dppLacZ*) in wild-type and lipophorin-RNAi discs stained in parallel and imaged under identical conditions. Discs from lipophorin-RNAi larvae activate *collier* at least as efficiently as those of the wild type (compare Fig. 4c and f). In contrast, the range of activation of *dppLacZ* is significantly narrowed in *lipophorin* RNAi discs. *dppLacZ* is expressed up to 11 cells away from the AP boundary in wild-type discs (Fig. 4a, b), but only up to six cells away in lipophorin-RNAi larvae (Fig. 4d, e, m). These data suggest that lipophorin knockdown decreases the range of Hedgehog signalling.

To discover whether Hedgehog trafficking was altered, we stained discs for Hedgehog and Patched. In wild-type discs, Hedgehog moves into the anterior compartment, where it is found in endosomes, often with Patched^{32,33} (Fig. 4g–i). Patched-mediated endocytosis is thought to sequester Hedgehog and limit its spread^{32,34}. Hedgehog is most abundant up to five cell rows away from the AP

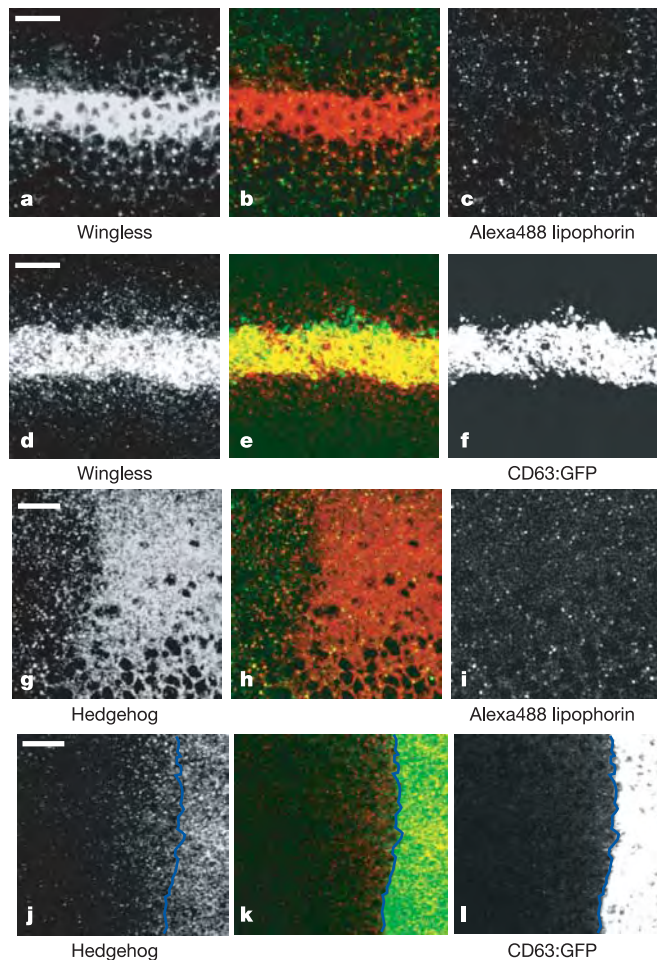


Figure 2 Wingless and Hedgehog co-localize with Alexa488 lipophorin. Scale bars = 10 μ m. Blue lines indicate AP boundaries. **a–c**, Disc stained with anti-Wingless (a, and red in b) after 20-min incubation with Alexa488 lipophorin (green in b, and c). **d–f**, Disc expressing CD63:GFP (green in b, and f) in Wingless-producing cells stained with anti-Wingless (d, and red in e). **g–i**, Disc stained with anti-Hedgehog (g, and red in h) after 20-min incubation with Alexa488 lipophorin (green in h, and i). **j–l**, Disc expressing CD63:GFP (green in k, and l) in Hedgehog-producing cells stained with anti-Hedgehog (j, and red in k).

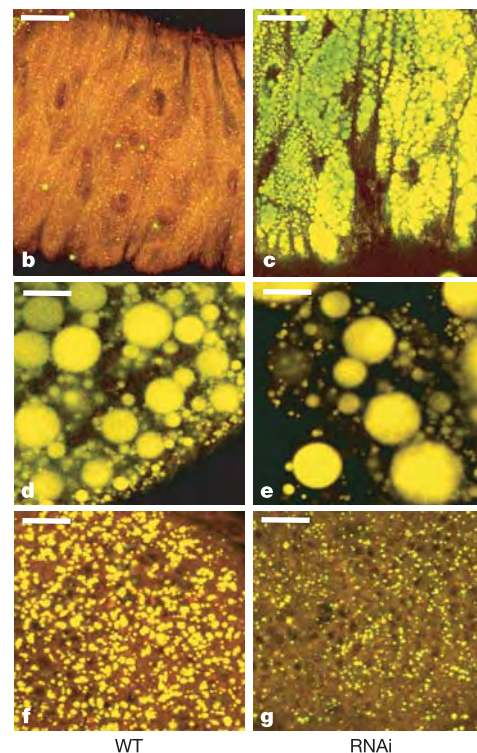


Figure 3 Lipophorin-RNAi perturbs lipid transport. **a**, Extracts from equal numbers of wild-type (WT) or *hs-flippase* +; *UAS dsRNA/Tubulin:GAL4* larvae probed for ApoLII at indicated hours after heat shock. **b–g**, Posterior midgut (b, c), fat body (d, e) or imaginal discs (f, g) from wild-type larvae (b, d, f) or *hs-flippase* +; *Adh:GAL4* +; *UAS dsRNA* + larvae (c, e, g) 5 days after heat shock. Yellow, neutral lipid; red, plasma membrane. Scale bar = 40 μ m (b, c) or 10 μ m (d–g).

boundary; although Hedgehog signals over a wider range, specific staining there cannot be distinguished from background. In lipophorin-RNAi discs, Hedgehog (Fig. 4j, k) accumulates to abnormally high levels in the first five rows of anterior cells. We counted 380 Hedgehog spots in the most apical 10 μm of the wild-type disc shown in Fig. 4g. The lipophorin-RNAi disc shown in Fig. 4j contains 1,208 Hedgehog spots in the same region. Most accumulated Hedgehog colocalizes with Patched (Fig. 4k, l) in endosomes (Supplementary Fig. S5). Furthermore, Patched co-accumulates more extensively with Hedgehog in endosomes than it does in wild-type (Fig. 4h, k). These data indicate that lipophorin RNAi either increases the susceptibility of Hedgehog to Patched-mediated endocytosis, or prevents subsequent degradation of the protein.

We wondered whether lipophorin depletion might affect Hedgehog trafficking indirectly by preventing release of a needed co-factor from some other larval tissue. To investigate this, we added purified lipophorin particles to explanted lipophorin-RNAi discs and examined Hedgehog and Patched distribution. Abnormal Hedgehog and Patched accumulation was strongly reduced by a two-hour incubation of dissected discs with lipophorin particles (Supplementary Fig. S7). Thus lipophorin acts directly in imaginal discs to control Hedgehog trafficking, although it is still possible that its effects on signalling are indirect.

Drosophila cannot synthesize sterols and relies on dietary sources. To assess whether reduced uptake of sterols or other lipids might cause the changes we see, we explored the effects of lipid deprivation on larval development. Larvae were allowed to hatch and feed on

sucrose/agarose plates supplemented with yeast for 2–3 days, then transferred to plates containing chloroform-extracted yeast autolysate, rather than yeast. These larvae are developmentally delayed; after 7 days of lipid deprivation, their discs are much smaller than those of younger late-third-instar larvae (compare Fig. 5a, b). In contrast, yeast-fed siblings pupariate and begin to eclose by this time. Those flies that infrequently eclose after larval lipid depletion are small (35–60% of normal body weight) but normally patterned (Fig. 5c, d). Thus, lipid depletion stalls imaginal growth.

To discover whether lipid starvation affected Hedgehog trafficking or signalling, we deprived larvae of lipid 2 days after hatching and stained their discs 6 days later (Fig. 5h–j). No changes in Hedgehog or Patched distribution are apparent in these discs compared with younger yeast-fed discs of similar size (Fig. 5e–g). Furthermore, the range of *dpp* and *collier* expression does not differ in lipid-starved and yeast-fed discs (Fig. 5k–p). Thus, lipid starvation does not mimic the effects of lipophorin knockdown. We speculate that lipid-starvation-induced growth arrest prevents membrane sterol from dropping to levels that would interfere with the Hedgehog pathway. Thus, lipophorin does not indirectly affect the Hedgehog pathway via lipid deprivation.

Wingless function requires lipophorin

To discover whether lipophorin RNAi perturbed Wingless trafficking, we examined Wingless distribution. In lipophorin-RNAi discs, extracellular Wingless is less abundant on both the apical and basolateral epithelial surfaces and spreads over shorter distances

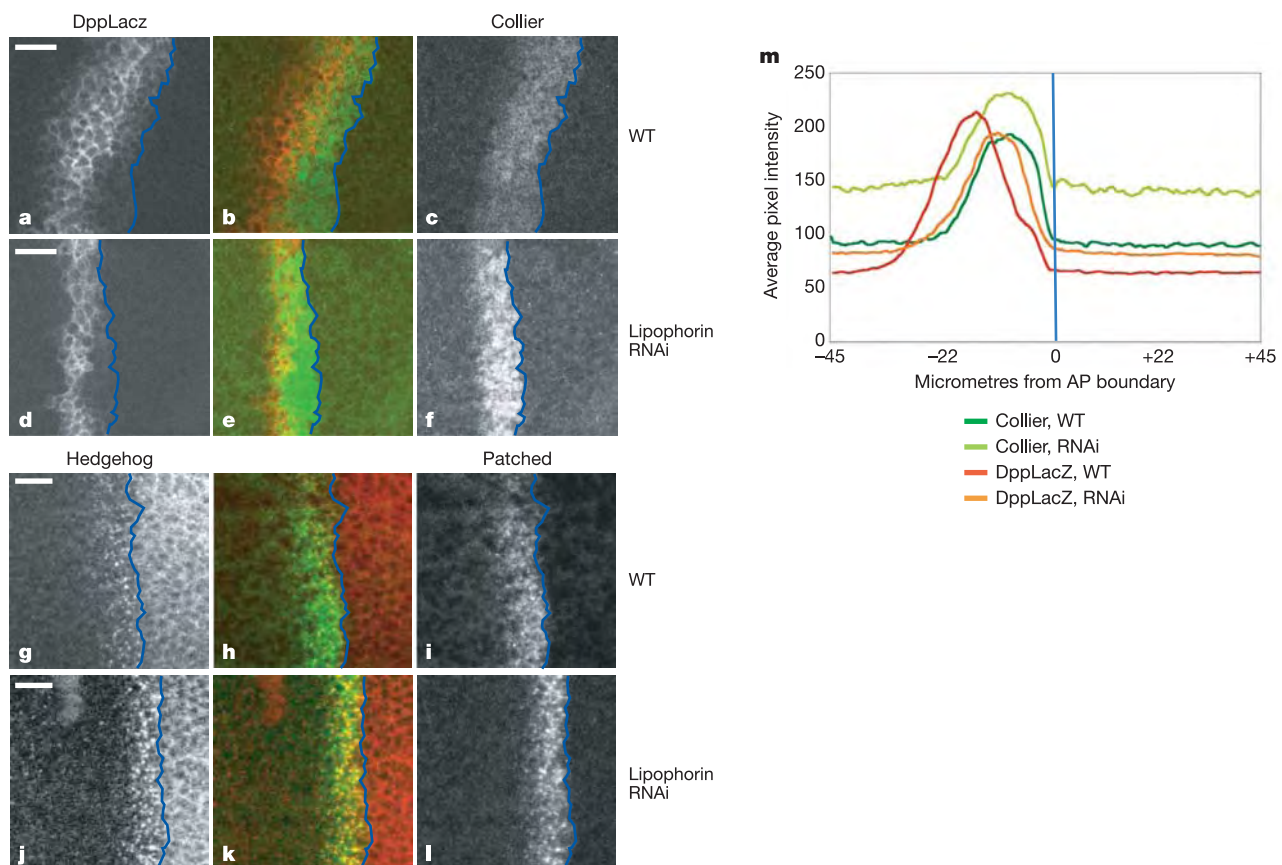


Figure 4 Lipophorin-RNAi alters Hedgehog distribution and signalling. Blue lines = AP boundary. Scale bar = 10 μm . **a–c**, *dpplacZ*+ disc 4 days after heat shock, stained for LacZ (**a**, and red in **b**) and Collier (green in **b**, and **c**). **d–f**, *hs-flippase*+/*dpplacZ*+/*Tubulin:GAL4/UAS:dsRNA* disc 4 days after heat shock, stained for LacZ (**d**, and red in **e**) and Collier (green in **e**, and **f**). **g–i**, Wild-type disc 4 days after heat shock, stained for Hedgehog (**g**, and red in **h**) and Patched (green in **h**, and **i**). **j–l**, *hs-flippase*

+/*UAS:dsRNA/Tubulin:GAL4* wing disc 4 days after heat shock, stained for Hedgehog (**j**, and red in **k**) and Patched (green in **k**, and **l**). **m**, Average Collier and DppLacZ staining intensities for four wild-type and four lipophorin-RNAi discs. Blue line indicates AP boundary. Average distance from AP boundary of peak LacZ staining was $16.6 \pm 2.7 \mu\text{m}$ for the wild type, and $11.1 \pm 1.5 \mu\text{m}$ for lipophorin-RNAi.

(Fig. 6a–d). However, no consistent alterations in intracellular Wingless are detected (not shown). Thus, lipophorin promotes accumulation of extracellular Wingless.

To investigate whether Wingless signalling required lipophorin, we examined the activation of two target genes. Senseless is produced only in cells near the Wingless source and its expression is unaffected by lipophorin RNAi (not shown). Distalless is normally produced in a gradient throughout most of the wing pouch. In lipophorin-RNAi discs, the Distalless gradient is abnormally narrow (Fig. 6e–g). This suggests that lipophorin knockdown specifically perturbs long-range Wingless signalling.

Here, we establish the principle that lipid-linked proteins of the exoplasmic face of the membrane associate with lipoproteins. These include many gpi-linked proteins with diverse functions, as well as the lipid-linked morphogens Wingless and Hedgehog. The mechanism allowing long-range dispersal of lipid-linked proteins is not yet understood. The finding that these proteins exist in both membrane-associated and lipoprotein-associated forms suggests reversible binding to lipoprotein particles as a plausible mechanism

for intercellular transfer, and the consequences of lowering lipoprotein levels in *Drosophila* larvae supports this idea.

Lipophorin knockdown narrows the range of both Wingless and Hedgehog signalling. Hedgehog accumulates to an abnormally high level in cells near the source of production and long-range signalling is inhibited; short-range target genes, however, are expressed normally. These data suggest that Hedgehog does not move as far when lipophorin levels are low. The range over which Hedgehog moves is normally restricted by Patched-mediated endocytosis. In discs from lipophorin RNAi larvae, accumulated Hedgehog co-localizes with Patched in endosomes, suggesting that it is more efficiently sequestered by Patched. How might lipophorin antagonize Patched-mediated sequestration and promote long-range movement?

Our data are consistent with the idea that lipophorin is continuously needed for movement, rather than required only for the release of morphogens. If lipophorin were important only for Hedgehog secretion, we would expect lipophorin RNAi to decrease the amount of Hedgehog found in receiving tissue; this seems not to

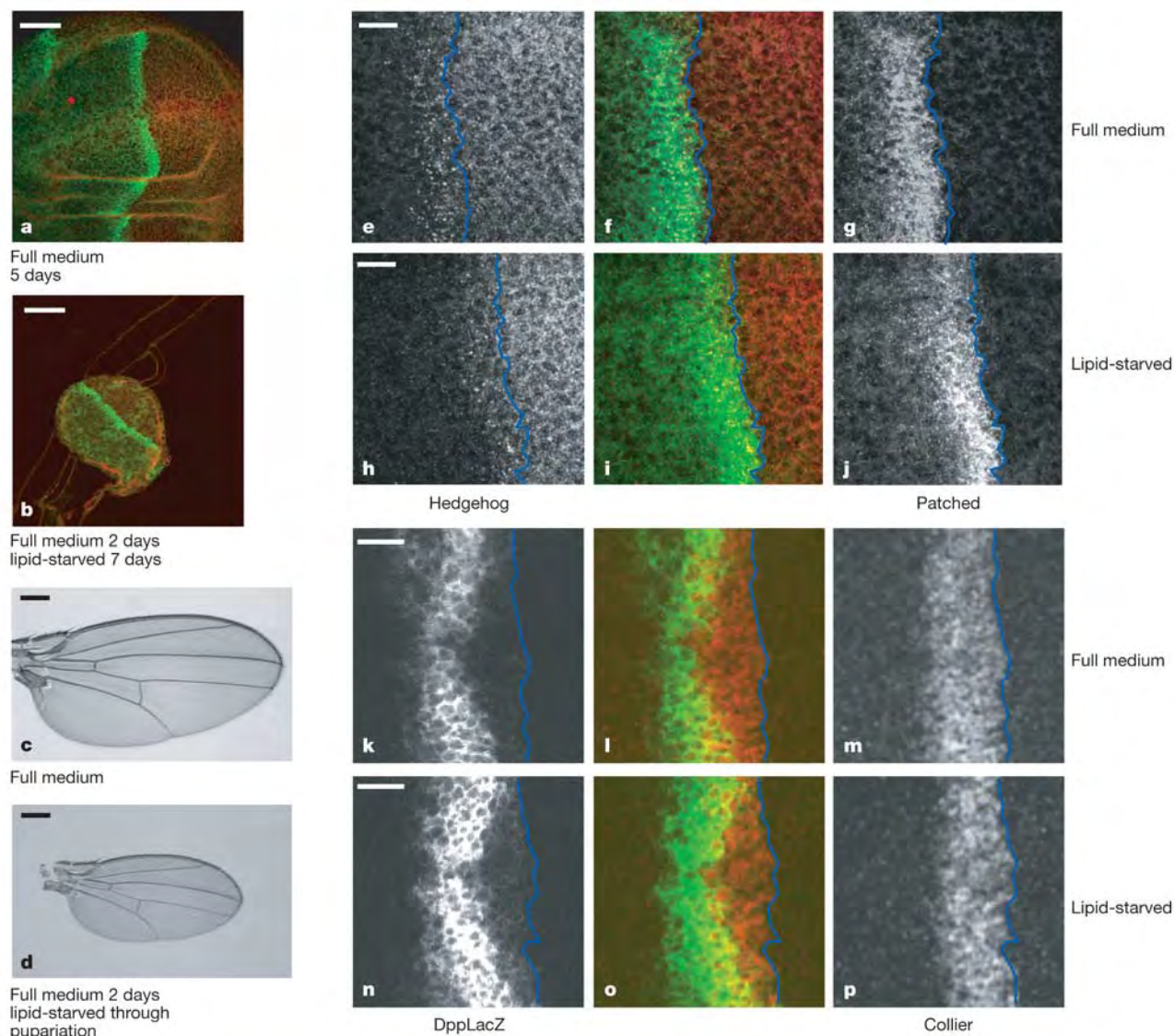


Figure 5 Hedgehog signalling is unaffected by lipid-depletion. **a, b**, Discs from fully fed (**a**) or lipid-starved (**b**) larvae. **c, d**, Wings from fully fed (**c**) or lipid-starved (**d**) flies. Scale bar = 250 μ m. **e–j**, Discs from fully fed (**e–g**) or lipid-starved (**h–j**) larvae stained for Hedgehog (**e, h**, and red in **f, i**) and Patched (green in **f, i**; and **g, j**). Scale bar = 30 μ m.

k–p, Disc from fully fed (**k–m**) or lipid-starved (**n–p**) *dpplacZ*⁺ larvae stained for LacZ (**k, n**, and green in **l, o**) and Collier (red in **l, o**; and **m, p**). Blue lines indicate AP boundary. Scale bars = 10 μ m.

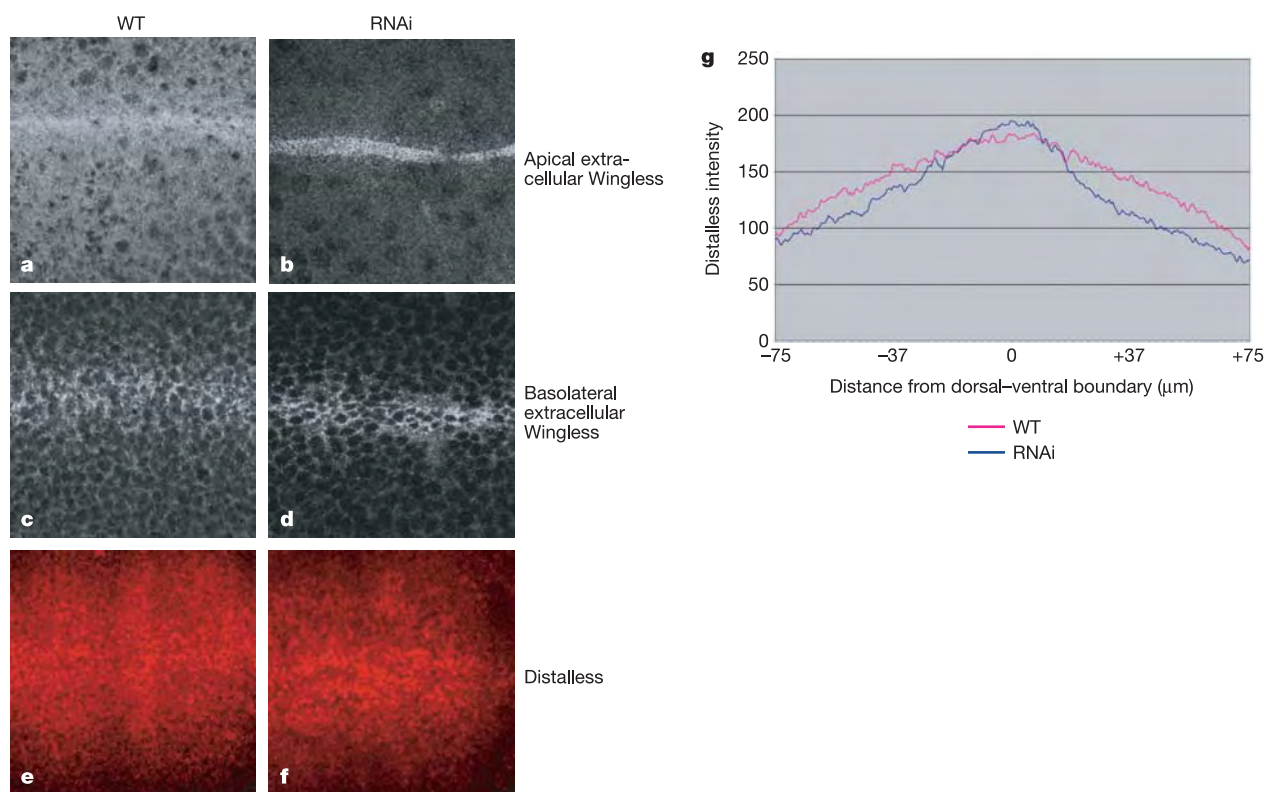


Figure 6 Lipophorin-RNAi narrows the range of Wingless signalling. **a–d**, Apical (**a**, **b**) and basolateral (**c**, **d**) sections of wild-type (**a**, **c**) and *Adh:GAL4/+; UASdsRNA/+* (**b**, **d**) wing discs 5 days after heat shock, stained for extracellular Wingless. **e**, **f**, Distalless protein accumulation in wild-type (**e**) and *hs-flippase/+; UAS dsRNA/ TubulinGAL4* (**f**)

wing discs 5 days after heat shock. **g**, Average Distalless staining intensity with distance from dorsal–ventral boundary of five wild-type (pink) and five *hs-flippase/+; UAS dsRNA/ TubulinGAL4* (blue) wing discs. For plots of individual discs, see Supplementary Fig. S6.

be the case. Furthermore, altered Hedgehog trafficking in receiving tissue is consistent with a model in which lipophorin is required at each step of intercellular transfer. We favour the idea that reversible association of Hedgehog with lipophorin particles facilitates its transfer from the plasma membrane of one cell to that of the next. This model predicts that lowering lipophorin levels should increase the length of time that Hedgehog spends in the plasma membrane before becoming associated with lipophorin. This would slow its rate of transfer and increase the probability of Patched endocytosing Hedgehog before it moved to the next cell. Hedgehog would then signal efficiently in the short range, but be so efficiently sequestered by Patched that very little protein would travel far enough to activate long-range target genes. These predictions are completely consistent with our observations.

This model differs significantly from our original concept of argosome function. We initially speculated that argosomes were exosome-like particles with an intact membrane bilayer, and that lipid-linked morphogens needed to be assembled on these particles to be secreted by producing cells. Instead, we find that argosomes are exogenously derived lipoproteins that facilitate the movement of morphogens through the epithelium. Many questions remain as to how morphogens become associated with argosomes, and how the spread and cell-interactions of these particles are regulated. Clearly, heparan sulphate proteoglycans are essential for the movement of Hedgehog and Wingless into receiving tissue^{35,36}. Because heparan sulphate binds to vertebrate lipoprotein particles^{37,38}, one might speculate that heparan sulphate proteoglycans (HSPGs) facilitate morphogen movement through lipoprotein binding. Conversely, we find many gpi-linked proteins, including the HSPG's Dally and Dally-like (unpublished data), on lipoprotein particles themselves. These associated proteins have the potential to modulate the cellular

affinities or trafficking properties of lipoproteins and the morphogens they carry.

Our data suggest that lipophorin particles not only mediate intercellular transfer of Hedgehog, but may also be endocytosed together with the morphogen. Interestingly, LDL-receptor-related proteins Arrow and Megalin have demonstrated roles in Wingless signalling and Hedgehog endocytosis, respectively^{39–41}. It is intriguing to speculate that these receptors might be important for interaction with the lipoprotein-associated form of the morphogen.

Cholesterol has the potential to modulate the activity of the Hedgehog pathway at many different points^{3,42–44}. Whether changes in the level of cellular cholesterol normally play a role in regulating the activity of the pathway is unclear. Here we show that Hedgehog interacts with the particle that delivers sterol to cells. This observation raises the possibility that internalization of Hedgehog is linked to sterol uptake, and suggests new mechanisms to link nutrition, growth and signalling during development. □

Methods

Fractionation

Five millilitres of larvae were homogenized with 5 ml of 150 mM NaCl, 50 mM Tris-Cl pH 7.4, 2 mM EGTA plus protease inhibitors on ice. The 1,000g supernatant was centrifuged for 3 h at 33,600 r.p.m. (120,000g) at 4 °C in a SW40Ti rotor generating a pellet (P120) and supernatant (S120).

For isopycnic density centrifugation, we added 0.33 g ml⁻¹ KBr to the S120, and centrifuged for 2 days at 40,000 r.p.m. (285,000g) at 10 °C in an SW40Ti rotor.

Immunoprecipitation

S120 pre-cleared 2 h with protein A-sephacryl CL4B beads was incubated with beads linked to different sera. Beads were washed with PBS, 1% BSA, then PBS, and eluted using Laemmli sample buffer or 150 mM NaCl, 2 mM EDTA, 100 mM Tris-Cl pH 8.3, 0.5% Nonidet-P40, 0.5% sodiumdeoxycholate and 0.1% SDS.

Antisera

Rabbits were immunized with synthetic peptide LEGVIRDSPKFKDL (Hedgehog amino acids 123–138), conjugated to keyhole limpet haemocyanin (Eurogentec). Antibody was affinity-purified on peptide-conjugated affigel-15 columns (Biorad).

DNA encoding amino acids 195–509 or 891–1070 (parts of ApoLII and ApoLI, respectively) was amplified from GH18004 (Resgen) and cloned into pQE30. His-tagged fusion proteins (Qiagen) were used to immunize rats or rabbits.

Expression construct

CD63:EGFP amplified from pEGFP-C1-bos (gift from G. Griffiths) was cloned into pUAST⁴⁵.

RNA interference

RNA interference was induced by expressing inverted repeats derived from two different regions of the Pro-apolipophorin cDNA (607 bp ending 47 bp from stop codon, and 500 bp starting at ATG). The first was amplified and inserted into pENTR2B (Invitrogen). Using the Gateway system, we inserted it twice in inverted orientation into pFRIPE. pFRIPE is derived from pUAST; downstream of the UAS are two Gateway insertion sites flanking an FLP cassette containing the HcRed gene and a transcription termination sequence.

The second fragment was amplified and cloned as an inverted repeat into pUhr. pUhr was derived from pUAST by inserting an HcRed-containing FLP cassette between the UAS and the multiple cloning site.

Flies expressing lipophorin–RNAi constructs were crossed with others harbouring heat-shock-inducible FLP and one of several GAL4 drivers. After 5 days at 25 °C, larvae were heat-shocked for 90 min at 37 °C; this causes excision in all cells as determined by HcRed fluorescence. No excision occurs without heat shock in any larval tissue except the fat body (not shown).

dsRNA expressed under the control of either TubulinGAL4 (ubiquitous), AdhGAL4 (fat body and part of the gut) or C76S GAL4 (disc-specific) was semi-lethal and produced identical larval phenotypes. No phenotype was ever observed when *lipophorin* dsRNA was expressed in imaginal discs.

Immunohistochemistry

Imaginal discs were fixed and stained as described¹³. Antibodies were diluted as follows: anti-Wg⁴⁶, 1:200; anti-Hh⁴⁷, 1:500; 1:100; anti-Ptc⁴⁸, 1:50; anti-βgal (Promega Z378A) 1:100; anti-Col³¹, 1:200. To compare wild type and lipophorin–RNAi animals, tissues were stained in parallel and imaged under identical conditions with an LSMZeiss or Leica confocal microscope.

Image analysis

Hedgehog-positive spots in wild-type and lipophorin–RNAi discs were quantified in ten confocal sections 1 μm apart. The signal threshold was adjusted to 130 and images were despeckled using ImageJ (<http://rsb.info.nih.gov/ij/>). Grids were overlaid on the processed image and spots were counted manually.

We used ImageJ to quantify the Hedgehog signalling range in five projected apical sections of Col- and Dpp-stained discs. For each image, we determined pixel intensity along ten lines centred at the AP boundary using the Plot Profile function of ImageJ and averaged them to obtain a plot for each disc. Average plots from four discs of each type were generated using Microsoft Excel. Distalless range in Fig. 6 and Patched staining in Supplementary Fig. S7 were quantified similarly.

Lipid starvation

Eggs were collected on apple juice/agar plates + yeast for 24 h, then allowed to develop for 2–3 days on the same yeast-containing plates. Larvae were rinsed with PBS + 0.05% TritonX100, treated for 10 s with 50% Na hypochlorite, and rinsed with sterile H₂O. Larvae were transferred with sterile forceps to 10-cm plates containing 2% chloroform extracted agarose, 2.5% sucrose and 0.15% Nipagen, supplemented with either 0.3 g chloroform extracted yeast autolysate (for lipid starvation), or 0.3 g yeast (for lipid-fed controls).

Labelling lipophorin with Alexa488

Lipophorin particles were fluorescently labelled with AlexaFluor 488 (Molecular Probes) according to the manufacturer's instructions. Conjugate was separated from un-reacted label using Sephadex G-25 PD-10 columns (Amersham Pharmacia Biotech) and eluted with 100-mM Na-phosphate, pH 7.4, 100 mM NaCl, 10% sucrose.

Incubation of dissected discs with lipophorin particles

For experiments shown in Fig. 2 and Supplementary Fig. 7, imaginal discs were incubated at 29 °C with 50 μg ml^{−1} lipophorin particles for 20 min and 2 h, respectively. On the basis of the starting volume of larvae and the final volume in which lipophorin was eluted, we estimate that this represents approximately 1/10 of the concentration of lipoprotein particles present in the haemolymph.

Received 10 November 2004; accepted 28 February 2005; doi:10.1038/nature03504.

- Strigini, M. & Cohen, S. M. A Hedgehog activity gradient contributes to AP axial patterning of the *Drosophila* wing. *Development* **124**, 4697–4705 (1997).
- Neumann, C. J. & Cohen, S. M. Long-range action of Wingless organizes the dorsal-ventral axis of the *Drosophila* wing. *Development* **124**, 861–870 (1997).

- Porter, J. A. *et al.* Hedgehog patterning activity: role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. *Cell* **86**, 21–34 (1996).
- Willert, K. *et al.* Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448–452 (2003).
- Pepinsky, R. B. *et al.* Identification of a palmitic acid-modified form of human Sonic hedgehog. *J. Biol. Chem.* **273**, 14037–14045 (1998).
- Williams, K. P. *et al.* Functional antagonists of sonic hedgehog reveal the importance of the N terminus for activity. *J. Cell Sci.* **112**, 4405–4414 (1999).
- Chamoun, Z. *et al.* Skinny hedgehog, an acyltransferase required for palmitoylation and activity of the hedgehog signal. *Science* **293**, 2080–2084 (2001).
- Michelli, C. A., The, I., Selva, E., Mogila, V. & Perrimon, N. Rasp, a putative transmembrane acyltransferase, is required for Hedgehog signaling. *Development* **129**, 843–851 (2002).
- Lee, J. D. *et al.* An acylatable residue of Hedgehog is differentially required in *Drosophila* and mouse limb development. *Dev. Biol.* **233**, 122–136 (2001).
- Liu, T. *et al.* Intercellular transfer of the cellular prion protein. *J. Biol. Chem.* **277**, 47671–47678 (2002).
- Dunn, D. E. *et al.* A knock-out model of paroxysmal nocturnal hemoglobinuria: Pig-a(−) hematopoiesis is reconstituted following intercellular transfer of GPI-anchored proteins. *Proc. Natl Acad. Sci. USA* **93**, 7938–7943 (1996).
- Anderson, S. M., Yu, G., Giattina, M. & Miller, J. L. Intercellular transfer of a glycosylphosphatidylinositol (GPI)-linked protein: release and uptake of CD4-GPI from recombinant adeno-associated virus-transduced HeLa cells. *Proc. Natl Acad. Sci. USA* **93**, 5894–5898 (1996).
- Greco, V., Hannus, M. & Eaton, S. Argosomes: a potential vehicle for the spread of morphogens through epithelia. *Cell* **106**, 633–645 (2001).
- Denzer, K., Kleijmeer, M. J., Heijnen, H. F., Stoorvogel, W. & Geuze, H. J. Exosome: from internal vesicle of the multivesicular body to intercellular signaling device. *J. Cell Sci.* **113**, 3365–3374 (2000).
- Arrese, E. L. *et al.* Lipid storage and mobilization in insects: current status and future directions. *Insect Biochem. Mol. Biol.* **31**, 7–17 (2001).
- van der Horst, D. J., van Hoof, D., van Marrewijk, W. J. & Rodenburg, K. W. Alternative lipid mobilization: the insect shuttle system. *Mol. Cell. Biochem.* **239**, 113–119 (2002).
- Escola, J. M. *et al.* Selective enrichment of tetraspan proteins on the internal vesicles of multivesicular endosomes and on exosomes secreted by human B-lymphocytes. *J. Biol. Chem.* **273**, 20121–20127 (1998).
- Wubbolts, R. *et al.* Proteomic and biochemical analyses of human B cell-derived exosomes. Potential implications for their function and multivesicular body formation. *J. Biol. Chem.* **278**, 10963–10972 (2003).
- Sundermeyer, K., Hendricks, J. K., Prasad, S. V. & Wells, M. A. The precursor protein of the structural apolipoproteins of lipophorin: cDNA and deduced amino acid sequence. *Insect Biochem. Mol. Biol.* **26**, 735–738 (1996).
- Kutty, R. K. *et al.* Molecular characterization and developmental expression of a retinoid- and fatty acid-binding glycoprotein from *Drosophila*. A putative lipophorin. *J. Biol. Chem.* **271**, 20641–20649 (1996).
- Hortsch, M. & Goodman, C. S. *Drosophila* fasciclin I, a neural cell adhesion molecule, has a phosphatidylinositol lipid membrane anchor that is developmentally regulated. *J. Biol. Chem.* **265**, 15104–15109 (1990).
- Nose, A., Mahajan, V. B. & Goodman, C. S. Connectin: a homophilic cell adhesion molecule expressed on a subset of muscles and the motoneurons that innervate them in *Drosophila*. *Cell* **70**, 553–567 (1992).
- Butler, S. J., Ray, S. & Hiromi, Y. klongin, a novel member of the *Drosophila* immunoglobulin superfamily, is required for the development of the R7 photoreceptor neuron. *Development* **124**, 781–792 (1997).
- Incardona, J. P. & Rosenberry, T. L. Replacement of the glycoinositol phospholipid anchor of *Drosophila* acetylcholinesterase with a transmembrane domain does not alter sorting in neurons and epithelia but in behavioral defects. *Mol. Biol. Cell* **7**, 613–630 (1996).
- Cagan, R. L., Kramer, H., Hart, A. C. & Zipursky, S. L. The bride of sevenless and sevenless interaction: Internalization of a transmembrane ligand. *Cell* **69**, 393–399 (1992).
- Klueg, K. M. & Muskavitch, M. A. Ligand-receptor interactions and trans-endocytosis of Delta, Serrate and Notch: members of the Notch signalling pathway in *Drosophila*. *J. Cell Sci.* **112**, 3289–3297 (1999).
- Culi, J. & Mann, R. S. Boca, an endoplasmic reticulum protein required for wingless signaling and trafficking of LDL receptor family members in *Drosophila*. *Cell* **112**, 343–354 (2003).
- Greenspan, P., Mayer, E. P. & Fowler, S. D. Nile red: a selective fluorescent stain for intracellular lipid droplets. *J. Cell Biol.* **100**, 965–973 (1985).
- Britton, J. S., Lockwood, W. K., Li, L., Cohen, S. M. & Edgar, B. A. *Drosophila*'s insulin/P13-kinase pathway coordinates cellular metabolism with nutritional conditions. *Dev. Cell* **2**, 239–249 (2002).
- Vervoort, M. hedgehog and wing development in *Drosophila*: a morphogen at work? *Bioessays* **22**, 460–468 (2000).
- Vervoort, M., Crozatier, M., Valle, D. & Vincent, A. The COE transcription factor Collier is a mediator of short-range Hedgehog-induced patterning of the *Drosophila* wing. *Curr. Biol.* **9**, 632–639 (1999).
- Torres, C., Gorfinkel, N. & Guerrero, I. Patched controls the Hedgehog gradient by endocytosis in a dynamin-dependent manner, but this internalization does not play a major role in signal transduction. *Development* **131**, 2395–2408 (2004).
- Tabata, T. & Kornberg, T. B. Hedgehog is a signaling protein with a key role in patterning *Drosophila* imaginal discs. *Cell* **76**, 89–102 (1994).
- Chen, Y. & Struhl, G. Dual roles for patched in sequestering and transducing Hedgehog. *Cell* **87**, 553–563 (1996).
- The, I., Bellaiche, Y. & Perrimon, N. Hedgehog movement is regulated through tout velu-dependent synthesis of a heparan sulfate proteoglycan. *Mol. Cell* **4**, 633–639 (1999).
- Takei, Y., Ozawa, Y., Sato, M., Watanabe, A. & Tabata, T. Three *Drosophila* EXT genes shape morphogen gradients through synthesis of heparan sulfate proteoglycans. *Development* **131**, 73–82 (2004).
- Mahley, R. W. & Ji, Z. S. Remnant lipoprotein metabolism: key pathways involving cell-surface heparan sulfate proteoglycans and apolipoprotein E. *J. Lipid Res.* **40**, 1–16 (1999).

38. Camejo, G., Hurt-Camejo, E., Wiklund, O. & Bondjers, G. Association of apo B lipoproteins with arterial proteoglycans: pathological significance and molecular basis. *Atherosclerosis* **139**, 205–222 (1998).
39. McCarthy, R. A., Barth, J. L., Chintalapudi, M. R., Knaak, C. & Argraves, W. S. Megalin functions as an endocytic sonic hedgehog receptor. *J. Biol. Chem.* **277**, 25660–25667 (2002).
40. Wehrli, M. *et al.* arrow encodes an LDL-receptor-related protein essential for Wingless signalling. *Nature* **407**, 527–530 (2000).
41. Tamai, K. *et al.* LDL-receptor-related proteins in Wnt signal transduction. *Nature* **407**, 530–535 (2000).
42. Burke, R. *et al.* Dispatched, a novel sterol-sensing domain protein dedicated to the release of cholesterol-modified hedgehog from signaling cells. *Cell* **99**, 803–815 (1999).
43. Nakano, Y. *et al.* A protein with several possible membrane-spanning domains encoded by the *Drosophila* segment polarity gene patched. *Nature* **341**, 508–513 (1989).
44. Cooper, M. K. *et al.* A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nature Genet.* **33**, 508–513 (2003).
45. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
46. Strigini, M. & Cohen, S. M. Wingless gradient formation in the *Drosophila* wing. *Curr. Biol.* **10**, 293–300 (2000).
47. Taylor, A. M., Nakano, Y., Mohler, J. & Ingham, P. W. Contrasting distributions of patched and hedgehog proteins in the *Drosophila* embryo. *Mech. Dev.* **42**, 89–96 (1993).
48. Capdevila, J., Pariente, F., Sanpedro, J., Alonso, J. & Guerrero, I. Subcellular localization of the segment

polarity protein Patched suggests an interaction with the Wingless reception complex in *Drosophila* embryos. *Development* **120**, 987–988 (1994).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank T. Kurzchalia for advice regarding lipid depletion. We thank S. Cohen, I. Guerrero, P. Ingham, Y. Hiromi, M. Horscht, John Incardona and R. White for gifts of antibodies, and G. Griffiths for the CD63:GFP fusion construct. We are grateful to A. Mahmoud and S. Bowman for developing CFP:Rab5-expressing flies, and to V. Greco for helping to initiate these studies. We thank D. Backasch for performing embryo injections. K. Simons, M. Zerial, T. Kurzchalia and C. Dahmann provided comments on the manuscript.

Author contributions This work has been a collaborative effort between the groups of C. Thiele and S. Eaton, the Thiele laboratory contributing biochemical expertise and the Eaton laboratory expertise in working with *Drosophila*. First authors appear in alphabetical order.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.E. (eaton@mpi-cbg.de) or C.T. (thiele@mpi-cbg.de).

Compositional maps of Saturn's moon Phoebe from imaging spectroscopy

Roger N. Clark¹, Robert H. Brown², Ralf Jaumann³, Dale P. Cruikshank⁴, Robert M. Nelson⁵, Bonnie J. Buratti⁵, Thomas B. McCord⁶, J. Lunine², K. H. Baines⁷, G. Bellucci⁷, J.-P. Bibring⁸, F. Capaccioni⁷, P. Cerroni⁷, A. Coradini⁷, V. Formisano⁷, Y. Langevin⁸, D. L. Matson⁵, V. Mennella⁷, P. D. Nicholson⁹, B. Sicardy¹⁰, C. Sotin¹¹, Todd M. Hoefen¹, John M. Curchin¹, Gary Hansen¹³, Karl Hibbitts¹² & K.-D. Matz³

¹US Geological Survey, MS964, Box 25046, Federal Center, Denver, Colorado 80225, USA

²Lunar and Planetary Laboratory and Stewart Observatory, University of Arizona, Tucson, Arizona 85721, USA

³German Aerospace Center (DLR), Institute of Space Sensor Technology and Planetary Exploration, Rutherfordstrasse 2, D-12489 Berlin, Germany

⁴NASA Ames Research Center, Moffett Field, California 94035, USA

⁵Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

⁶University of Hawaii at Manoa, HIGP/SOEST, 1680 East-West Road, Honolulu, Hawaii 96822, USA

⁷Istituto di Fisica dello Spazio Interplanetario, CNR, Rome 00133, Italy

⁸Institut d'Astrophysique Spatiale, Université de Paris-Sud, F-91405 Orsay Cedex France

⁹Cornell University, Astronomy Department, Ithaca, New York 14853, USA

¹⁰Observatoire de Paris, 92195 Meudon, France

¹¹Laboratoire de Planétologie et Géodynamique, UMR CNRS 6112, Université de Nantes, 44322 Nantes, France

¹²Planetary Science Institute NW, Corporate Center Pasadena, 255 S. Lake Avenue, Suite 300, Pasadena, California 91101, USA

¹³Department of Earth and Space Sciences, University of Seattle, Washington 98195-1310, USA

The origin of Phoebe, which is the outermost large satellite of Saturn, is of particular interest because its inclined, retrograde orbit suggests that it was gravitationally captured by Saturn, having accreted outside the region of the solar nebula in which Saturn formed¹. By contrast, Saturn's regular satellites (with prograde, low-inclination, circular orbits) probably accreted within the sub-nebula in which Saturn itself formed². Here we report imaging spectroscopy of Phoebe resulting from the Cassini-Huygens spacecraft encounter on 11 June 2004. We mapped ferrous-iron-bearing minerals, bound water, trapped CO₂, probable phyllosilicates, organics, nitriles and cyanide compounds. Detection of these compounds on Phoebe makes it one of the most compositionally diverse objects yet observed in our Solar System. It is likely that Phoebe's surface contains primitive materials from the outer Solar System, indicating a surface of cometary origin.

We report the identification and possible origins of the surface materials on Phoebe by compositional mapping with the Visible and Infrared Mapping Spectrometer³ (VIMS) on board Cassini. Cassini VIMS spatially resolved the surface of Phoebe, and data have been registered and projected into simple cylindrical maps: low to medium resolution with near hemispheric coverage, and high resolution of a small area (Fig. 1). Spectra of Phoebe indicate a low albedo surface, from <1% to 6% reflectance (Fig. 2), with a variety of absorption features. These features are due to materials (Table 1) that occur with variable abundances and/or grain sizes in different locations on the body. They include: previously identified water ice⁴, bound water, and trapped CO₂ (maps, Fig. 1; spectra, Fig. 2). A broad 1-μm feature, attributed to Fe²⁺-bearing minerals, is present in almost all regions but stronger in equatorial areas. Bound water and hydroxyl absorptions indicate the presence of phyllosilicates. Other weak spectral

features may be caused by cyanide and organic compounds.

The composition of Phoebe should reflect the composition of the region in the solar nebula where it accreted. If Phoebe originated in the region of the asteroid main belt (~3 AU), it should consist largely of volatile-poor mafic minerals. Alternatively, it could have accreted in the volatile-rich outer solar nebula where the Kuiper belt objects (KBOs) originated. Dynamical scattering inward from an accretion site among the KBOs is more likely than scattering outward from the asteroid belt⁵. For example, inward scattering may have occurred during events that are thought to have forced the KBOs outward by the outward migration of Neptune⁶.

With a diameter of 220 km, Phoebe is comparable in size to most of the nearly 800 known KBOs. Other physical properties that Phoebe shares with KBOs include the presence of H₂O ice and an overall low surface albedo⁷. KBOs exhibit a wide range of colour⁸, including the neutral colour characteristic of Phoebe⁹. It has been postulated that material from Phoebe (for example, ejected by impacts) may coat the leading surface of Saturn's moon Iapetus¹⁰, so this material might be found elsewhere in the Saturn system.

Mapping results from VIMS indicate that H₂O ice is distributed over most of Phoebe's observed surface, but generally shows stronger spectral signatures toward the southern polar region (Fig. 1b, h). In the highest spatial resolution data, crater interiors display less exposed ice than the surrounding terrain (arrows in Fig. 1b). In the jovian system, cratering tends to expose fresh ice in the subsurface¹¹. The highest resolution VIMS data for Phoebe also indicate an ice-rich layer exposed in crater walls just below Phoebe's surface (Fig. 1b). Therefore, it is likely that ice is less abundant in the deeper subsurface, at least in the region of highest resolution data (near longitude 0°, latitude 0°).

The Fe²⁺ absorption is quite broad (Fig. 2b, feature f1), unlike that in pyroxenes, but similar to that in olivines and phyllosilicates¹². If phyllosilicates were present, (OH)⁻ absorptions might be expected, but would be relatively weak owing to the low reflectance. Indeed, weak 2.2- and 2.3-μm absorptions (f10, f23) are

Table 1 Observed spectral features

Feature	Wavelength (μm)	Width (μm)	Origin
f1	1.0	1.1	Probable Fe ²⁺
f2	1.04	~0.05	H ₂ O ice overtone
f3	1.2	~0.02	OH stretch, CH combination, or calibration artefact
f4	1.3	~0.02	OH stretch, CH combination, or calibration artefact
f5	1.4	~0.02	OH stretch, CH combination, or calibration artefact
f6	1.5	~0.2	H ₂ O ice overtone
f7	1.7	Variable	CH stretch overtone
f8	1.95	~0.1	Bound H ₂ O
f9	2.02	~0.2	H ₂ O ice combination
f10	2.16	~0.03	Probable metal-OH combination, for example, phyllosilicates
f11	2.42	~0.07	Probable CN combination
f12	2.95	~0.7	H ₂ O ice and/or bound H ₂ O
f13	3.1	~0.05	H ₂ O ice Fresnel peak
f14	3.2-3.3	Variable	CH stretch fundamentals
f15	3.55	~0.06	Probable CH
f16	3.9	~0.2	Perhaps CH, CN or H ₂ O
f17	4.26	~0.03	Trapped CO ₂ (gaseous or fluid inclusion)
f18	4.50	~0.03	CN in a nitrile
f19	4.8-5.0	~0.05	Probable CN fundamental
f20	>5.1	ND	ND
f21	4.3	0.7	H ₂ O ice and/or bound water
f22	2.05	~0.17	H ₂ O ice
f23	2.3	~0.08	Probable metal-OH combination, for example, phyllosilicates
f24	2.72	~0.05	OH stretch fundamental, common in phyllosilicates
f25	1.25	~0.1	H ₂ O ice
f26	3.62	~0.07	Probable CH or CN
f27	2.54	~0.016	Probable CH

ND, not determined.

seen in some spectra, indicating a probable metal-OH absorption (Fig. 2b). Hydrated mineral absorptions are also consistent with phyllosilicates. We attribute the 1.5- and 1.95- μm absorptions (Fig. 2b, spectrum S2, f6, f8) to hydrated mineral absorptions, but these features are not diagnostic of specific mineralogy and indicate only the presence of bound water.

CO_2 is ubiquitous on Phoebe. The wavelength and shape of the CO_2 absorption, 4.26 μm (Fig. 2a, c, d, f17), is similar to that of the trapped CO_2 found on Jupiter's satellites Callisto and Ganymede¹¹. CO_2 is more spatially variable than water ice in the areas where it maps on Phoebe's surface (Fig. 1g). The highest spatial resolution data indicate a linear structure of lower CO_2 that is at least 60 km in length (Fig. 1g). Medium resolution data indicate other linear

structures of variable CO_2 abundances, showing structures only seen in composition maps (Fig. 1g) and not observed in reflectance (Fig. 1a).

The absorption feature at 2.42 μm (Fig. 2a, c, e, f11) has not been reported on any Solar System body before Cassini. Searches of available spectral libraries^{12,13} and scientific papers that show OH-bearing minerals with an absorption at this position also have other absorptions nearby that are not seen in Phoebe spectra. Cyanide compounds, however, have unique absorptions in this region, and a spectrum of potassium ferrocyanide trihydrate, $\text{K}_4\text{Fe}(\text{CN})_6 \cdot 3\text{H}_2\text{O}$, closely matches Phoebe's absorption position and shape (Fig. 2e). Spatially, the 2.42- μm absorption maps mainly in the equatorial regions (Fig. 1e). The same 2.42- μm absorption has been observed

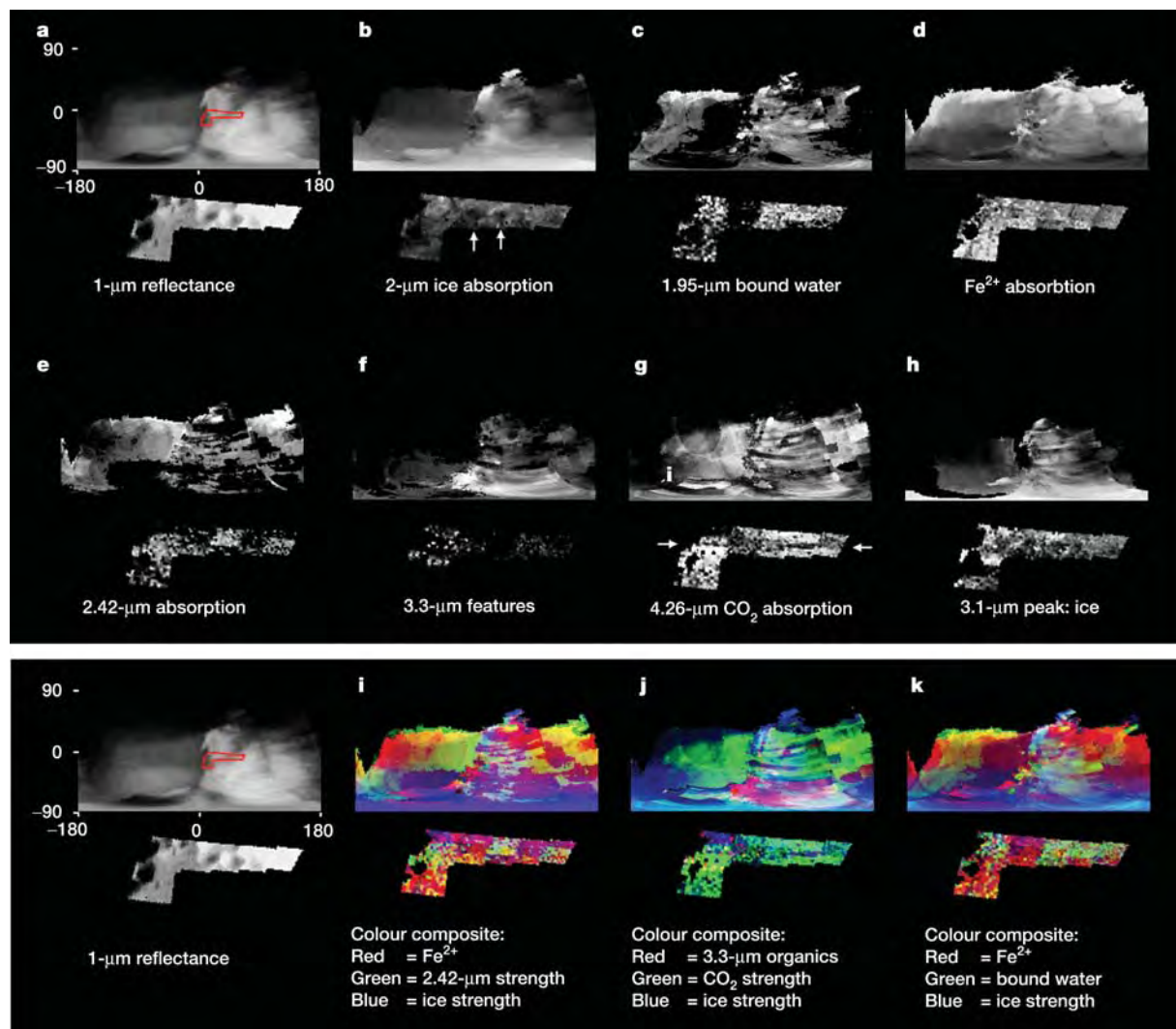


Figure 1 VIMS data, x and y axes show longitude and latitude, respectively, in degrees. **a–h**, Compositional maps of Phoebe in simple cylindrical projection, VIMS obtained spatially resolved hyperspectral images of Phoebe at 352 wavelengths (0.4–5 μm) from 8:47 UT June 11, 2004 at an initial range of 245,833 km and phase angle of 84.9°, to 10:22 UT June 12, 2004 at a final range of 338,401 km and phase angle of 92.2°. The closest image was obtained on 19:32 UT June 11 at a range of 2,178 km and solar phase angle of 24.6°. For each map, two resolutions are shown: medium resolution (top) sampled at 0.5° per pixel (about 1.9 km at the equator), and high resolution sampled at 0.48 km pixel⁻¹ (original VIMS resolution was 0.96 km pixel⁻¹). The spatial mapping of the VIMS, with an instantaneous field of view of 0.25 × 0.5 mrad, resulted in spatial coverage at full spectral resolution (~0.016 μm) up to 1 km per pixel, allowing many diagnostic absorption features to be mapped. **a**, The apparent reflectance at 1 μm is shown, ignoring lighting due to slope variations. The red outline shows the region of the

high resolution image. **b–h**, The continuum-removed absorption band depth is shown for each material, as defined in ref. 21. **b**, The 2- μm ice absorption strength. Note that two craters (arrows) have no ice in the floors. The crater on the right shows enhanced ice strength near the top of the crater (bright arc). Other materials map in the floors of the two craters, for example, Fe^{2+} -minerals (**c**). Maps of probable cyanide compounds (**e**), probable organics (**f**), CO_2 (**g**), and high abundance crystalline water ice (**h**) are shown. A curious linear structure seen in the map of CO_2 is indicated by the two arrows in **g**. That structure is not seen in other compositional maps, nor in reflectance (**a**). **i–k**, Colour composite mineral maps showing complex spatial relationships of materials across Phoebe's surface, compared to the reflectance image (**a**; shown again as the black and white image on left). **i**, Fe^{2+} -bearing minerals, probable cyanide compounds, and water ice. **j**, Organic materials, CO_2 , and water ice. **k**, Fe^{2+} -bearing minerals, bound water and water ice.

in the first VIMS data of the dark side of Iapetus¹⁴. A similar absorption was also seen in comet Borrelly¹⁵.

Certain regions of Phoebe exhibit weak absorptions in the 3.3- μm region (Figs 1f and 2d, f14), indicating the presence of organics. Both the position of the CH stretch fundamental near 3.3 μm , and the lack of absorptions in the 3.4–3.7 μm region¹⁶, argue for an aromatic or cycloalkane hydrocarbon origin and against long chain

aliphatics. The lack of spectral structure observed in the VIMS spectral range for pyrene and anthracene rejects these aromatics.

Other areas of Phoebe show absorption features at 4.50 and 2.54 μm (f18 and f27), as well as in the 3.3- μm region (Figs 1f and 2a, d). The narrow 4.5- μm absorption indicates a CN triple bond ($\text{C}\equiv\text{N}$). The best solution for these spectra appears to be an alkane molecule (single bonded CC) with an attached CN, a nitrile group.

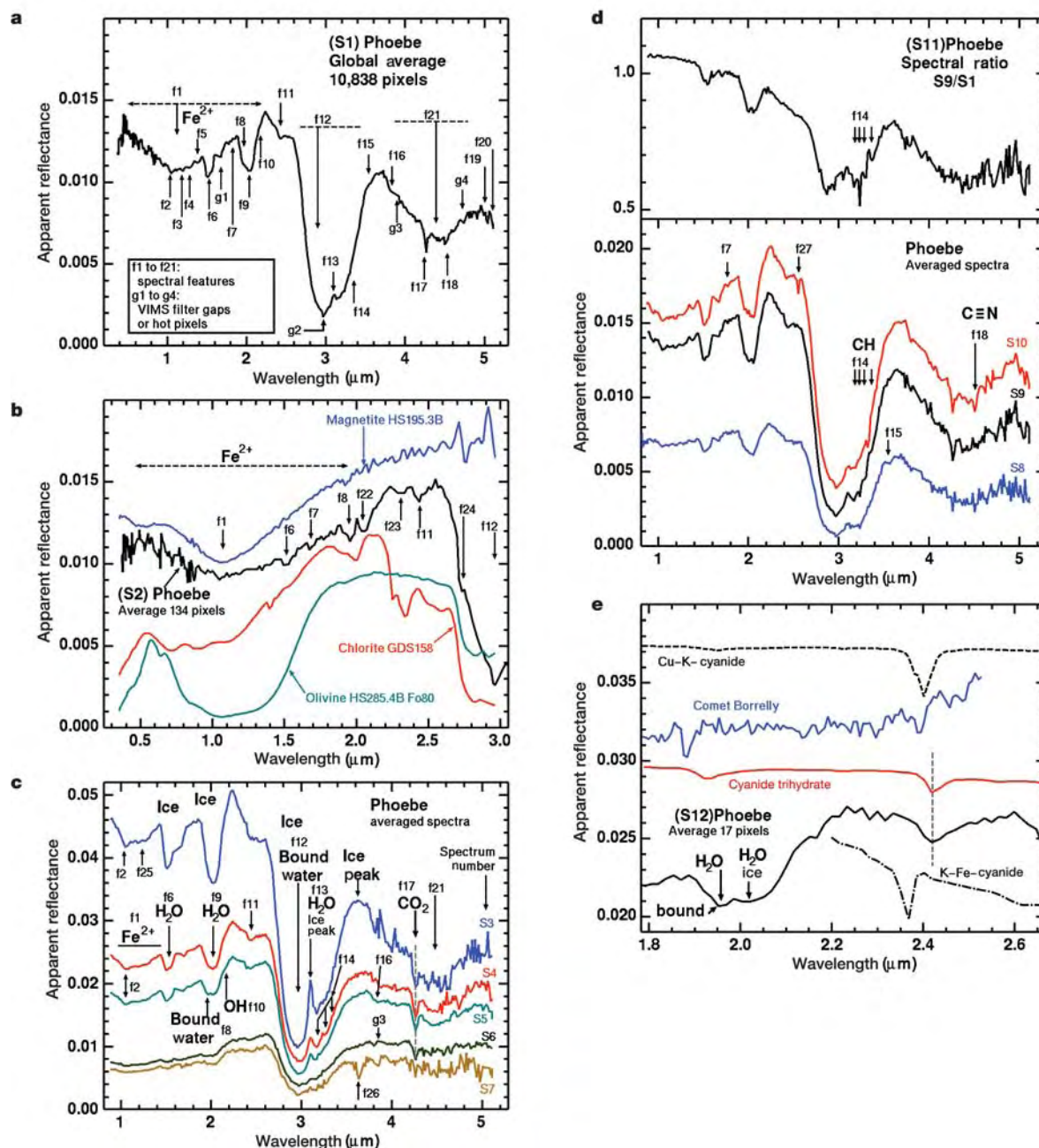


Figure 2 Spectra of Phoebe show the presence of numerous compounds. Spectra are labelled S1–S12, and spectral features are labelled f1–f27. VIMS grating filter gaps are labelled g1–g4, and should be ignored. **a**, An average of over 10,000 VIMS pixels shows the highest signal-to-noise ratio spectrum, covering most of the images shown here (avoiding high emission angle pixels near the edge of the map). The CO_2 absorption (f17) occurs at a wavelength shorter than that observed in spectra of solid CO_2 . Other spectra (**b–e**) are averages from the brightest pixels in Fig. 1, excluding high emission angle pixels near the edge of the map. **b**, Phoebe spectra compared to three of many Fe^{2+} -bearing minerals¹² with similar absorptions. **c**, Illustration of the spectral diversity of Phoebe's surface. These spectra are averages from multiple small locations on Phoebe's surface, obtained from unprojected data. **d**, Phoebe spectra showing organic and nitrile

absorptions. **e**, Comparison of Phoebe spectra to cyanide compounds¹² and to spectra of comet Borrelly¹⁵. Cyanide compounds have a CN fundamental stretch in the 4.7–5- μm region, and the position of the fundamental as well as the 2.4- μm absorption varies with the compound's specific composition¹². Although the VIMS' signal-to-noise ratio in that spectral region is low, areas on Phoebe with the strongest 2.42- μm absorption show spectral structure similar to that expected from a cyanide origin. Because the CN fundamental would show saturated spectral structure, the observed spectral structure would be weak. The broad 3.9- and 4.5- μm features (**a**, **c**, f16, f21) may be organic related, or due to spectral structures in the Fe-bearing minerals, bound water-bearing minerals, water ice, or other materials not directly observed in the spectra.

Our studies indicate that organic materials with a few per cent nitrile approximately match the Phoebe spectra. A spectral feature at 3.62 μm is seen in some Phoebe spectra (Fig. 2c, f26, Table 1). This absorption is not accompanied by other nearby features that might indicate an organic origin. A spectrum of HCN (ref. 16) shows a similar isolated absorption.

The low albedo of Phoebe is probably caused by carbon and/or organic molecules. Some asteroids with low albedo have been associated with certain meteorite classes that are rich in organic materials^{17,18}. Unfortunately, the lower spectral resolution of the asteroid data prohibits direct comparison to the VIMS Phoebe spectra.

Spectra of Phoebe display a wealth of information, indicating a surface containing distinct locations of ferrous-iron-bearing minerals (Fig. 1d), bound water (Fig. 1c), trapped CO_2 (Fig. 1g), probable phyllosilicates, organics (Fig. 1f), nitriles and cyanide compounds (Fig. 1e). The only body imaged to date that is more compositionally diverse is Earth. Phoebe's organic and cyanide compositions are unlike any surface yet observed in the inner Solar System. Although this composition is more similar to comets, the water ice may only be a surface coating because some crater interiors show less water ice deep in the crater and more water near the surface (Fig. 1b). This raises the possibility that Phoebe is coated by material of cometary or outer Solar System origin. Without information about its deep internal composition, it is not possible to conclude that Phoebe has an outer Solar System origin. Nevertheless, compositional data for the full Saturn system may help constrain Phoebe's origin. The same broad Fe^{2+} absorption on Phoebe is seen in Saturn's rings, particularly in the Cassini division and the C-ring¹⁹, and may imply that some materials are common to both Phoebe's surface and the rings. But organics and cyanide compounds have yet to be definitively detected in any VIMS Saturn ring data to date, despite much higher signal-to-noise ratios than obtained on Phoebe. This may mean that the materials on Phoebe and the rings have different origins.

The low albedo and neutral colour (at visible wavelengths) of Phoebe are probably caused by the presence of carbon, iron-bearing minerals, and some quantity of complex organic compounds. Organic molecules tend to impart red coloration to a planetary surface, although a moderate degree of colour can also come from mafic minerals such as Mg-rich pyroxene²⁰. The neutral reflectance of Phoebe suggests that amorphous or moderately structured elemental carbon dominates the colour, and that the effects of the minerals and organic complexes are minimal. The detection of compounds with a 2.42- μm absorption on both Phoebe and Iapetus may indicate that material from Phoebe has struck Iapetus's leading hemisphere. Alternatively, perhaps cometary material has coated both Phoebe and Iapetus. Regardless of its origin, Phoebe's diverse mix of surface materials is unique among Solar System surfaces observed to date, and it probably samples primitive materials in the outer Solar System. □

Received 21 December 2004; accepted 11 March 2005; doi:10.1038/nature03558.

- Pollack, J. B. *et al.* Gas drag in primordial circumplanetary envelopes: A mechanism for satellite capture. *Icarus* **37**, 587–611 (1979).
- Burns, J. A. in *Satellites* (eds Burns, J. A. & Matthews, M. S.) 117–158 (Univ. Arizona Press, Tucson, 1986).
- Brown, R. H. *et al.* The Cassini Huygens Mission. *Space Sci. Rev.* (in the press).
- Owen, T. C. *et al.* Detection of water ice on Saturn's satellite Phoebe. *Icarus* **139**, 379–382 (1999).
- Liou, J. C. & Malhotra, R. Depletion of the outer asteroid belt. *Science* **275**, 375–377 (1997).
- Levison, H. F. & Morbidelli, A. The formation of the Kuiper belt by the outward transport of bodies during Neptune's migration. *Nature* **426**, 419–421 (2003).
- Brown, R. H., Cruikshank, D. P. & Pendleton, Y. J. Water ice on Kuiper Belt object 1996 TO₆₆. *Astrophys. J. Lett.* **519**, 101–104 (1999).
- Barucci, M. A. *et al.* Analysis of Trans-Neptunian and Centaur colours: continuous trend or grouping? *Astron. Astrophys.* **371**, 1150–1154 (2001).
- Buratti, B. J., Hicks, M. D., Tryka, K. A., Sittig, M. S. & Newburn, R. L. High-resolution 0.33–0.92 μm spectra of Iapetus, Hyperion, Phoebe, Rhea, Dione, and D-type asteroids: How are they related? *Icarus* **155**, 375–381 (2002).
- Jarvis, K. S., Vilas, F., Larson, S. M. & Gaffey, M. J. Are Hyperion and Phoebe linked to Iapetus? *Icarus* **146**, 125–132 (2000).

- McCord, T. B. *et al.* Non-water-ice constituents in the surface material of the icy Galilean satellites from Galileo Near-Infrared Mapping Spectrometer investigation. *J. Geophys. Res.* **103**, 8603–8626 (1998).
- Clark, R. N. *et al.* The U.S. Geological Survey, Digital Spectral Library splib05 (Open File Report 03–395, USGS, Denver, 2003); (<http://speclab.cr.usgs.gov/spectral-lib.html>).
- Hook, S. *ASTER Spectral Library* (Jet Propulsion Laboratory, Pasadena, 2004); (<http://speclib.jpl.nasa.gov>).
- Buratti, B. J. *et al.* Iapetus: First data from the Cassini Visual Infrared Mapping Spectrometer. *Bull. Am. Astron. Soc.* **36**, 1072 (2004).
- Soderblom, L. A. *et al.* Observations of comet 19P/Borrelly by the Miniature Integrated Camera and Spectrometer aboard Deep Space 1. *Science* **296**, 1087–1091 (2002).
- Stein, S. E. NIST Standard Reference Database 35. *NIST Spectral Library* (National Institute of Standards and Technology, Gaithersburg, Maryland, 2004); (<http://www.nist.gov/srd/nist35.htm>).
- Gaffey, M. J. Forging an asteroid-meteorite link. *Science* **260**, 167–168 (1993).
- Cruikshank, D. P. *et al.* Search for 3.4- μm C-H spectral bands on low albedo asteroids. *Icarus* **156**, 434–441 (2002).
- Brown, R. H. *et al.* Cassini's Visual and Infrared Mapping Spectrometer (VIMS): Observations during approach and orbit insertion. *Astrophys. J. Lett.* (in the press).
- Cruikshank, D. P. *et al.* Constraints on the composition of Trojan asteroid 624 Hektor. *Icarus* **153**, 348–360 (2001).
- Clark, R. N. *et al.* Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res.* **108**, 5131, doi:10.1029/2002JE001847 (2003).

Acknowledgements This work was funded by the Cassini project. Authors from American institutions were funded by NASA; authors from European institutions were funded by ESA.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.N.C. (rclark@usgs.gov).

Saturn's moon Phoebe as a captured body from the outer Solar System

Torrence V. Johnson¹ & Jonathan I. Lunine²

¹Jet Propulsion Laboratory, California Institute of Technology, M/S 264-525 4800 Oak Grove Drive, Pasadena, California 91109, USA

²Lunar and Planetary Laboratory, 1629 East University Boulevard, Tucson, Arizona 85721, USA

The orbital properties of Phoebe, one of Saturn's irregular moons, suggest that it was captured by the ringed planet's gravitational field rather than formed *in situ*. Phoebe's generally dark surface shows evidence of water ice¹, but otherwise the surface most closely resembles that of C-type asteroids² and small outer Solar System bodies such as Chiron and Pholus that are thought to have originated in the Kuiper belt³. A close fly-by of Phoebe by the Cassini–Huygens spacecraft on 11 June 2004 (19 days before the spacecraft entered orbit around Saturn) provided an opportunity to test the hypothesis that this moon did not form *in situ* during Saturn's formation, but is instead a product of the larger protoplanetary disk or 'solar nebula'. Here we derive the rock-to-ice ratio of Phoebe using its density^{4,5} combined with newly measured oxygen and carbon abundances in the solar photosphere^{6,7}. Phoebe's composition is close to that derived for other solar nebula bodies such as Triton and Pluto, but is very different from that of the regular satellites of Saturn, supporting Phoebe's origin as a captured body from the outer Solar System.

One characteristic that might help determine whether Phoebe originated in the solar nebula or the 'circum-saturnian' disk is the ratio of water ice to other materials in its interior, inferred from the average density of the object. The mean density of the 'regular' satellite system of Saturn is 1,300 kg m^{-3} , with an uncertainty of approximately $\pm 10\%$ (ref. 8). This is the mass-weighted value for Mimas, Enceladus, Tethys, Dione, Rhea and Iapetus—moons with prograde, in-plane, essentially circular orbits well inward of

Phoebe's retrograde and eccentric orbit. The uncertainty is based on those reported for the individual satellite densities, weighted by their mass. We excluded irregular Hyperion because it may be a remnant of a larger object disrupted by collision. This moon's mass is so small that its exclusion does not alter the system's mean density to the level of accuracy we give here. We also excluded Titan. Its density, corrected for the strong effects of compression of the ice caused by its large mass, is $1,600 \text{ kg m}^{-3}$, but Titan is massive enough that loss of water during formation might have altered significantly its overall ratio of rock to ice. Phoebe's density, based on the mass determined by measuring the perturbation of the spacecraft's trajectory associated with the Phoebe fly-by and the volume determined by analysis of imaging science subsystem data, is $1,630 \pm 33 \text{ kg m}^{-3}$ (refs 4, 5).

Phoebe's mass is too small to compress water ice or silicate significantly, so the derived density for Phoebe is equal to that of the material in the absence of pressure effects. Likewise, the central pressure of Phoebe ($\sim 4 \text{ MPa}$) is more than an order of magnitude less than that required for high-pressure phases of water ice to be stable. However, the low pressures in the interior of this body also allow for significant void space, or porosity, to be maintained. Porosities of small Solar System bodies are poorly constrained, but have been estimated for various small asteroids and satellites as being between 0.2 and 0.6 (ref. 9). Figure 1 shows the range of Phoebe's equivalent 'sample density' (without porosity) for various assumptions for porosity. If Phoebe's porosity were 0.5, it could be made entirely of anhydrous silicates. However, Cassini's Visible and Infrared Mapping Spectrometer (VIMS) detected abundant water ice and water bound to minerals¹⁰, making a purely silicate composition unlikely. We therefore take the plausible range of porosity of Phoebe to be from near zero to about 50%.

The uncompressed densities of Pluto, a Kuiper belt object, and Neptune's satellite Triton, which may have originated in the Kuiper belt before being captured by Neptune, are both around $1,900 \text{ kg m}^{-3}$ (refs 11, 12). If Phoebe were derived from the same compositional reservoir as Pluto and Triton, its present porosity would have to be ~ 0.15 to attain the same material density. Even

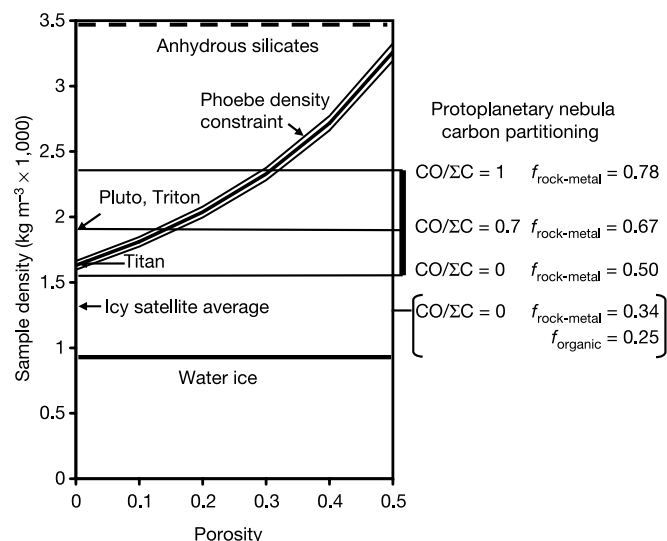


Figure 1 Plot of the material or 'sample' density versus porosity of Phoebe using data from the Cassini orbiter. The heavy line is the central reported value, $1,630 \text{ kg m}^{-3}$, with lighter lines on either side representing the uncertainty of $\pm 33 \text{ kg m}^{-3}$. Densities of uncompressed water ice and anhydrous silicates are shown for comparison. Also labelled are the densities of Pluto and Triton, the mass-weighted mean density for the 'regular' icy saturnian satellites, and the density of Titan. The calculated sample densities discussed in the text are shown for different ratios of carbon contained in CO to total carbon: $\text{CO}/\Sigma\text{C} = \text{CO}/(\text{CO} + \text{CH}_4)$.

if the porosity of Phoebe were zero, its Cassini-derived density would be 1σ above that of the regular icy saturnian satellites. Therefore, Phoebe appears to be compositionally different from the intermediate-sized regular satellites of Saturn.

To quantify the compositional differences, we calculated the mass fractions of rock and ice in Phoebe, constrained by its observed density and the range of plausible porosities. For a protoplanetary disk of cosmic composition the dominant solid-forming element is oxygen (O), which will combine with silicon and magnesium to form magnesium silicates (calcium-aluminium silicates are less abundant because of the smaller cosmic abundances of Al and Ca compared with Mg), with iron to form various oxides, and with hydrogen to form water¹³. We adjusted the rock-to-ice ratio by taking solar abundances of magnesium, silicon and oxygen and forming MgSiO_3 (enstatite) as the 'rock' phase, which is seen in chondrites. We assume that sulphur takes up iron, but as the latter is more abundant we allow FeO to form until the iron is depleted, and add nickel to this 'metal' phase. The nickel abundance is sufficiently small to have little bearing on the density. The oxygen remaining is partitioned into water and carbon-bearing molecules, mostly carbon monoxide.

The abundance of CO that was in the protoplanetary or circum-saturnian disk is poorly constrained; a very reducing disk would have essentially all the carbon in methane (CH_4) and other organic phases. Carbon dioxide is not important as a sink for oxygen in any protoplanetary or circum-saturnian disk model¹³. We set $\text{CO}/(\text{CO} + \text{CH}_4)$ to be a free parameter and examined the consequences of allowing this to range from 0 to 1. We then computed the mass fractions of rock (f_{rock}), metal (f_{metal}) and water ice (f_{ice}) via¹⁴:

$$\rho_{\text{Phoebe}} = [(f_{\text{rock}}/\rho_{\text{rock}}) + (f_{\text{metal}}/\rho_{\text{metal}}) + (f_{\text{ice}}/\rho_{\text{ice}})]^{-1}$$

We adopted standard densities for the silicate ($3,360 \text{ kg m}^{-3}$), which was assumed to be anhydrous and is consistent with the density of the almost-certainly dehydrated moon Io, the metallic sulphide/oxide phase ($4,800 \text{ kg m}^{-3}$), and water ice (940 kg m^{-3})¹⁵. The hydration state of the silicates in the protoplanetary disk, or circum-saturnian disk, is poorly constrained. Most cosmochemists argue that hydration was limited in the former¹⁶, but Jupiter's moon Europa has an ice layer whose mass is consistent with dehydration of the moon's silicate bulk¹⁷. However, if we choose hydrated silicates in our model—and hence a lower density for the silicate phase—the amount of water declines commensurately. Therefore, there is little difference in making one choice or the other, and for concreteness we consider anhydrous silicates only.

Because the carbon abundance is approximately half that of oxygen in a solar composition gas, the precise abundances of these two elements are crucial in quantifying the rock mass fraction in Phoebe and the other saturnian satellites. Recent reanalysis of solar photospheric abundances^{6,7} have revised the oxygen abundance, and to a much lesser extent the carbon abundance, downward relative to the previously accepted values¹⁸. Expressed in the standard way, with the logarithmic (base 10) abundance of hydrogen set to 12.00, oxygen is 8.66 and carbon is 8.39. For these values, and currently accepted silicate/metal abundances¹⁸, we find the rock and metal mass fraction $f_{\text{rock-metal}} = 0.785$ and $\rho_{\text{Phoebe}} = 2,310 \text{ kg m}^{-3}$ when CO is the sole carbon-bearing molecule. When CH_4 is the only carbon-bearing molecule, $f_{\text{rock-metal}} = 0.505$ and $\rho_{\text{Phoebe}} = 1,520 \text{ kg m}^{-3}$. The upper density is consistent with Phoebe's measured density if the satellite has a porosity of ~ 0.3 . The lower density is not consistent with the measured density. However, allowing CO to be 25% of the carbon budget (with CH_4 the rest) yields a density of $1,620 \text{ kg m}^{-3}$ (and $f_{\text{rock-metal}} = 0.555$), consistent with Phoebe's density for a porosity of zero.

Thus, as shown in Fig. 1, the density of Phoebe can be reproduced with the latest solar abundances of the elements and reasonable

porosities for a wide range of ratios of CO (oxidized carbon) to CH₄ (reduced carbon). The upper (CO-dominated) density is of particular interest because it is close to but above the density of Pluto and Triton. A material density for Phoebe of 1,900 kg m⁻³, the uncompressed value for Pluto and Triton given in Fig. 1, requires CO to have been 70% rather than 100% of the total carbon budget where Phoebe formed. This is still consistent with the general view that CO dominated over CH₄ in the protoplanetary disk¹⁹, yet is also consistent with the observation of methane on Pluto and Triton²⁰. The corresponding porosity for Phoebe is then 15%. So a reasonable set of numbers gives Phoebe a composition identical to that of the only Kuiper belt objects for which we possess densities.

The new elemental abundances have been challenged as an artefact of partial gravitational settling in the Sun, where the abundances are measured, such that the true primordial abundances of O and C should be raised by 0.07 dex (ref. 21). Were we to adopt the resulting O and C values, a Phoebe with zero porosity would be reproduced in a nebular model with CO as 50% of the carbon budget, rather than 25%. For a porous Phoebe, a material density of 1,900 kg m⁻³ (identical to that of Pluto and Triton) is reproduced, with CO representing 85% of the nebular carbon budget. Neither of these scenarios represents a qualitative change from the discussion given above. In contrast, the old elemental abundances¹⁸ cannot reproduce the high densities of Phoebe, Pluto or Triton, even if CO represents 100% of the carbon budget in the solar nebula.

With the new (or modified new) C and O abundances, it is more difficult to explain the significantly lower densities of the regular saturnian satellites, which are all assumed to have zero porosity because of their much larger masses compared with Phoebe (although outer layers with low porosity have been suggested for some of the smallest²²). The mass-averaged value of 1,300 kg m⁻³ cannot be reproduced even if the carbon budget includes only CH₄ and no CO. It is worth noting that this was not as much of an issue with the old solar abundances¹⁸ used in studies before the recent revisions^{6,7}, as the CH₄-dominated case resulted in a model density close to 1,300 kg m⁻³. It is possible that the formation of Saturn somehow led to a more oxygen-rich or water-rich circum-saturnian disk relative to the protoplanetary disk²³.

Alternatively, the mean density can be reduced by retaining the solar elemental oxygen abundance but postulating a phase of condensed, refractory carbon present in the interiors of these satellites. This phase would bring carbon into the satellites as its own solid phase (in contrast to the volatile CO and CH₄, which would have had to be trapped in water ice, and would not have significantly changed the satellite bulk density), and hence lower the mass fraction of high-density rock and ice. Such refractory organics are plausible, but their identity is not constrained. Data from Cassini VIMS indicate the presence of hydrocarbon-type organics in the spectrum of Phoebe itself. If they were present in significant abundance, we would calculate a lower material density of Phoebe for any given choice of porosity. For example, if 25% (by mass) of the carbon in the circum-saturnian disk were in the form of a refractory hydrocarbon with a mean molecular weight of 100 and density of 1,000 kg m⁻³, the observed mean density of the regular saturnian satellites would be lowered to 1,300 kg m⁻³ relative to the value obtained if all the carbon were in the form of methane. Such a hydrocarbon could consist of polymers of acetylene, a common product of methane chemistry in the absence of liquid water. If the environment in which Phoebe formed had 5% of the carbon in the form of this hydrocarbon—with the rest being CO—then the predicted density of the moon for zero porosity would fall from 2,310 kg m⁻³ to 1,920 kg m⁻³, equal to the uncompressed values for Pluto and Triton. Future Cassini observations of the regular saturnian satellites may better characterize the existence and nature of surface refractory organics, particularly on Iapetus, whose dark component might well be or include a refractory organic phase. □

Received 24 December 2004; accepted 19 January 2005; doi:10.1038/nature03384.

1. Brown, M. Near-infrared spectroscopy of Centaurs and irregular satellites. *Astron. J.* **119**, 977–983 (2000).
2. Simonelli, D. P. et al. Phoebe: Albedo map and photometric properties. *Icarus* **138**, 249–258 (1999).
3. Cruikshank, D. P. et al. The composition of Centaur 5145 Pholus. *Icarus* **135**, 389–407 (1998).
4. Porco, C. et al. Observations of the Saturn system from the imaging science subsystem on Cassini. *Science* (in the press).
5. Jacobson, R. A. et al. The orbits of the major Saturnian satellites and the gravity field of the Saturnian system. *Proc. Am. Astron. Soc. DPS Meet.* **36**, 15.02 (2004).
6. Asplund, M. et al. Line formation in solar granulation. IV. [OI], O I and OH lines and the photospheric O abundance. *Astron. Astrophys.* **417**, 751–768 (2004).
7. Allende Prieto, C., Lambert, D. L. & Asplund, M. A reappraisal of the solar photospheric C/O ratio. *Astrophys. J.* **573**, L137–L140 (2002).
8. Jacobson, R. A. The orbits of the major Saturnian satellites and the gravity field of Saturn from spacecraft and Earth-based observations. *Astron. J.* **128**, 492–501 (2004).
9. Merline, W. J. et al. Discovery of a moon orbiting the asteroid 45 Eugenia. *Nature* **401**, 565–568 (1999).
10. Clark, R. N. et al. Compositional maps of Saturn's moon Phoebe from imaging spectroscopy. *Nature* doi:10.1038/nature03558 (this issue).
11. Stern, S. A., McKinnon, W. B. & Lunine, J. I. in *Pluto and Charon* (eds Stern, S. A. & Tholen, D. J.) 605–663 (Univ. Arizona Press, Tucson, 1997).
12. McKinnon, W. B., Lunine, J. I. & Banfield, D. in *Neptune and Triton* (ed. Cruikshank, D. P.) 807–877 (Univ. Arizona Press, Tucson, 1996).
13. Prinn, R. G. & Fegley, B. in *Origin and Evolution of Planetary and Satellite Atmospheres* (eds Atreya, S. K., Pollack, J. B. & Matthews, M. S.) 78–136 (Univ. Arizona Press, Tucson, 1989).
14. Simonelli, D. P. et al. The carbon budget in the outer solar system. *Icarus* **82**, 1–35 (1989).
15. Consolmagno, G. J. & Schaeffer, M. *Worlds Apart: A Textbook in Planetary Sciences* (Prentice-Hall, Englewood Cliffs, 1994).
16. Fegley, B. Kinetics of gas-grain reactions in the solar nebula. *Space Sci. Rev.* **92**, 177–200 (2000).
17. Ransford, G. A., Finnerty, A. A. & Collerson, K. D. Europa's petrological thermal history. *Nature* **289**, 21–24 (1981).
18. Anders, E. & Grevesse, N. Abundances of the elements: Meteoritic and solar. *Geochim. Cosmochim. Acta* **53**, 197–214 (1989).
19. Prinn, R. G. in *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. I.) 1005–1028 (Univ. Arizona Press, Tucson, 1993).
20. Owen, T. C. et al. Surface ices and the atmospheric composition of Pluto. *Science* **261**, 745–748 (1993).
21. Lodders, K. Solar system abundances and condensation temperatures of the elements. *Astrophys. J.* **591**, 1220–1247 (2003).
22. Eluszkiewicz, J. Compaction and internal structure of Mimas. *Icarus* **84**, 215–225 (1990).
23. Mosqueira, I. & Estrada, P. R. Formation of the regular satellites of giant planets in an extended gaseous nebula II: satellite migration and survival. *Icarus* **163**, 232–255 (2003).

Acknowledgements This study was supported by the Cassini Project. A portion of the work was done at the Jet Propulsion Laboratory, California Institute of Technology, under a contract from NASA.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.I.L. (jlunine@lpl.arizona.edu).

Manipulating spin and charge in magnetic semiconductors using superconducting vortices

Mona Berciu¹, Tatiana G. Rappoport^{2,3} & Boldizsár Jankó^{2,3}

¹Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

²Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA

³Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

The continuous need for miniaturization and increase in device speed¹ drives the electronics industry to explore new avenues of information processing. One possibility is to use electron spin to store, manipulate and carry information². All such 'spintronics' applications are faced with formidable challenges in finding fast and efficient ways to create, transport, detect, control and manipulate spin textures and currents. Here we show how

most of these operations can be performed in a relatively simple manner in a hybrid system consisting of a superconducting film and a paramagnetic diluted magnetic semiconductor (DMS) quantum well. Our proposal is based on the observation that the inhomogeneous magnetic fields of the superconducting film create local spin and charge textures in the DMS quantum well, leading to a variety of effects such as Bloch oscillations and an unusual quantum Hall effect. We exploit recent progress in manipulating magnetic flux bundles (vortices) in superconductors^{3,4} and show how these can create, manipulate and control the spin textures in DMSs.

The magnetic properties of superconductors (SCs) and DMSs are vastly different. When placed in a magnetic field, a type-II SC allows the field to penetrate in a non-uniform fashion, as Abrikosov vortices—the field lines are bundled up into flux quanta surrounded by superfluid eddy currents⁵. The length scale characterizing the field inhomogeneity is the decay length of the superfluid eddy currents surrounding the core of an isolated vortex. This is known as the penetration depth, λ , and is a few tens of nanometres for several SCs (for example, $\lambda = 39$ nm in Nb). In contrast, magnetic impurities (typically Mn) embedded in a paramagnetic DMS (such as $\text{Ga}_{1-x}\text{Mn}_x\text{As}$, $x \approx 1.5\text{--}5\%$) align their spins along the external magnetic field $\mathbf{B}(\mathbf{r})$, producing a local magnetization $\mathbf{M}(\mathbf{r}) \propto \mathbf{B}(\mathbf{r})$. The strong exchange $J(\mathbf{r})$ between the spins of these impurities and the spins σ of free charge carriers present in the DMS induces a local spin polarization of the charge carriers, $\int d\mathbf{r}/r \mathbf{M}(\mathbf{r}) \cdot \sigma \propto \mathbf{B}(\mathbf{r}) \cdot \sigma$. This is incorporated into a formally unchanged Zeeman energy, $E_Z = -\frac{1}{2}g_{\text{eff}}\mu_B\sigma \cdot \mathbf{B}(\mathbf{r})$, but now with an enormous effective gyromagnetic ratio $g_{\text{eff}} \approx 10^2$. The end result is a giant Zeeman splitting between charge carrier spin states aligned parallel and antiparallel to the magnetic field⁶.

A novel situation arises when these two materials are brought into close proximity (see Fig. 1a). The inhomogeneous magnetic field of an SC vortex creates a Zeeman potential landscape in the DMS quantum well. If a large enough field variation occurs on small length scales, magnetic field induced localization of charge carriers occurs in the DMS⁷. These localized states exhibit spin textures, which reflect the local magnetic field distribution. The charge and spin texture remains attached to and adiabatically follows a moving vortex, as long as the Doppler shift in the bound state energy is smaller than the binding energy. Thus, the vortex acts as a spin and charge tweezer: control of the vortices' locations and dynamics translates into controlled manipulation of the spin and charge textures in the DMS. A further advantage is that these effects are present in the paramagnetic DMS, and therefore a low Curie temperature is not a hindering factor. The devices, however, must operate below the superconducting transition temperature.

To illustrate these concepts, we first consider an isolated vortex in an SC-DMS bilayer structure. This limit is relevant for low fields when vortices are far apart, or when the DMS is covered by nanoscale SC dots with lateral size $d \approx \lambda$. The magnetic field⁸ of a vortex out of a half-space isotropic SC is used to compute numerically the energies E_n and wavefunctions $\psi_n(r)$ of the bound states localized by the vortex field in the DMS quantum well. For simplicity, we use DMS/SC parameters typical of $\text{Ga}_{1-x}\text{Mn}_x\text{As}/\text{Nb}$. We envisage the quantum well as being made from several DMS digital layers⁹, which are less disordered than bulk-doped DMS and have the ideal geometry. Good mobilities have been demonstrated in similar quantum wells¹⁰. Moreover, the charge carrier density can be controlled independently in digital layers. Nb is a canonical type-II SC. It can be grown via sputtering at room temperature—this limits Mn diffusion, preserving the quality of the quantum well and preventing magnetic doping of the SC. (Further details on material issues and the numerical solution are given in Supplementary Information). The results for the local density of states (LDOS) $|\Psi_\sigma(r, E_f)|^2$, or charge texture, are shown in Fig. 1. For any energy E_f , $|\Psi_\sigma(r, E_f)|^2 = \sum_n |\psi_n(r, \sigma)|^2 \delta(E_n - E_f)$. (We make the usual

assumption that the local tunnelling amplitude is a weak function of the energy.)

The charge distribution bound by the vortex is made predominantly of spin-up contributions, as seen from comparing Fig. 1b and c: whereas both spin components extend over the same area (roughly the size of the vortex), their maxima differ by about a factor of 10. The spin-down components appear because of the finite radial component of the vortex field (Fig. 1d). The spin-up components are concentrated right under the vortex core where the perpendicular field is largest, whereas the spin-down components are pushed away from the vortex core, close to where the radial magnetic field is largest. The total charge and spin distributions plotted in Fig. 1e show that the spin texture is strongly polarized along the axis of the vortex. These textures should be observable with spin-polarized scanning tunnelling spectroscopy. However, as discussed below, magneto-transport may provide the most direct signature of the appearance of these textures in the DMS.

The location of the vortices can be constrained spatially using nano-engineered grooves in the SC layer: the vortices are trapped in the regions where the SC film is thinnest. The position of the vortex in the groove can be controlled by sending a current through the SC layer. The current exerts a transversal force, which moves the vortex³, together with the texture it created in the DMS. The spin textures attached to single vortices can therefore be used as operational units, or as building blocks in systems with multiple vortices. A simple device controlling the distance between two spin textures is sketched in Fig. 2a, b. Other examples, ranging from

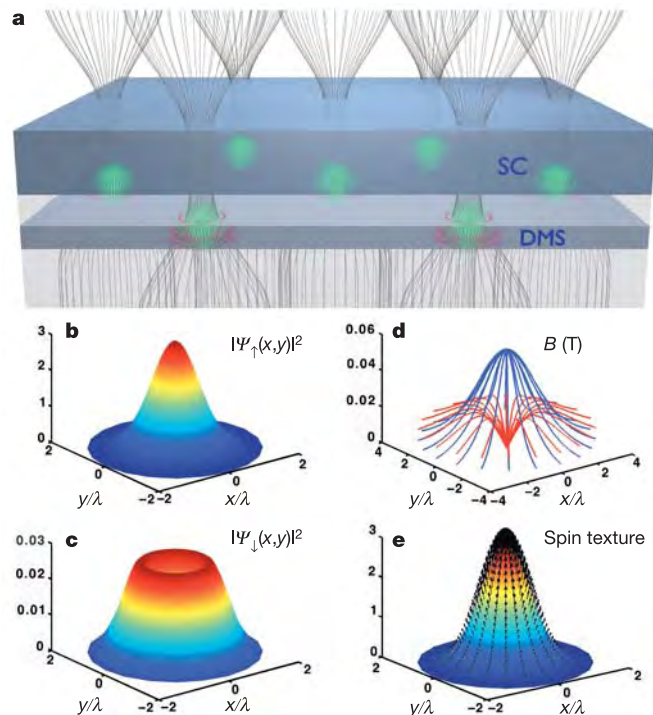


Figure 1 Spin and charge textures created by SC vortices in a DMS. **a**, Sketch of the hybrid structure, containing an SC layer in the vortex phase and a DMS quantum well. The field inhomogeneity of the vortex state is imprinted onto the DMS layer. Spin and charge textures are trapped in the high-field regions inside the DMS: **b**, charge texture for spin-up states; **c**, charge texture for spin-down states (note the smaller scale); **d**, spatial variation of the axial (blue) and radial (red) components of the isolated vortex field; **e**, total spin-charge texture. Arrows indicate the direction and relative magnitude of the local spin in the texture. In **b–e**, the DMS-SC distance is $z = 10$ nm and we use typical values of $\lambda \approx 40$ nm and a coherence length $\xi \approx 35$ nm for the SC, and an effective mass $m = 0.5m_e$ and a moderate $g_{\text{eff}} = 500$ for the DMS charge carriers.

using ratchets to create vortex (and thus spin) lenses and pumps, to quantum cellular automata, are described in Supplementary Information.

In uniform SC films, there is a simpler way to control the location of the vortices. The repulsion between vortices leads to the appearance of regular vortex lattices. As each vortex carries a magnetic flux $\phi = \phi_0/2$ (where $\phi_0 = h/e$ is the quantum of magnetic flux, h is Planck's constant, and e is the electronic charge), the distance between vortices can be continuously tuned with the external magnetic field. We first analyse the quasi-one-dimensional (1D) case, with long stripes of superconducting film, of width $w \approx \lambda$, deposited above a DMS quantum well. Such SC structures allow formation of 1D vortex chains. We use the 1D case to give an intuitive illustration of the various phenomena expected to occur in a tunable SC-DMS hybrid system, but we present numerical results only for the more complex 2D case. The low field limit (Fig. 2c) has the spin textures of a 1D set of isolated vortices. If the external magnetic field is increased, the inter-vortex distance $a(B)$ is reduced, together with the modulation of the Zeeman potential which traps the charges under individual vortices. Consequently, the bound states under each vortex are broadened into bands (Fig. 2d), and the trapped charges undergo a delocalization transition: they are now free to propagate along the 1D array. The new eigenstates of the charge carriers in the DMS are Bloch states of a 1D crystal, reflecting its translational symmetry. If the vortex array is along the x axis, then $\psi_k(x, y, z) = e^{ikx} u_k(x, y, z)$ with $u_k(x, y, z) = u_k(x + a(B), y, z)$ (see Supplementary Information for details). An immediate consequence of—or, conversely, the most convincing evidence for—the translational symmetry and the establishment of Bloch eigenstates is the appearance of the so-called Bloch oscillations: if the DMS is subjected to an electric field $E = E\hat{x}$, then an oscillating current should appear, with a characteristic period¹¹ $T = h/[eEa(B)]$. Bloch oscillations have never been observed in normal metals, owing to scattering off impurities before the Bloch states traverse the entire k -space Brillouin zone of size $2\pi/a(B)$. They have been detected only in artificial heterostructures where the lattice constant a is engineered to be much larger than in crystals. The structure that we suggest is unique, because in this case the periodicity $a(B) \approx B^{-1}$ is

tunable with the external magnetic field, and therefore Bloch oscillations can be searched for more easily.

The 2D case is more interesting. For an external magnetic field $B > B_{c1}$, the SC vortices order on a triangular lattice⁵. As each unit cell encloses the magnetic flux $\phi = Ba^2\sqrt{3}/2 = \phi_0/2$ of a vortex, the lattice constant $a \approx 1/\sqrt{B}$ is controlled by B . The 2D electron gas in the DMS quantum well is now in a periodic Zeeman potential and a perpendicular magnetic field. Its energy spectrum has a fractal-like structure as a function of ϕ/ϕ_0 (named the Hofstadter butterfly because of its shape). As $\phi/\phi_0 = q/p = 1/2$, the energy bands are those of a $q = 1, p = 2$ triangular Hofstadter butterfly^{12,13}. In the limit $a(B) \gg \lambda$, the vortices are well separated. The 'atom-like' states bound to each isolated vortex widen into energy bands, typical of an ordered crystal. The width of the bands is defined by the hopping amplitude t between neighbouring bound states. If t is small compared to the energy spacing between consecutive bound states, this Hofstadter problem corresponds to the regime of a dominant periodic modulation, which can be treated within a simple tight-binding model¹², and each band splits into p (here $p = 2$) magnetic sub-bands. This band structure has unique signatures in magnetotransport, as discussed below. Its measurement would provide a clear signature of the Hofstadter butterfly, which is a matter of considerable experimental interest^{14–16}. As the external field B is increased such that $a \approx \lambda$, the magnetic fields of different vortices start to overlap significantly. In this limit, the total magnetic field in the DMS layer is almost homogeneous. Its average is $B\hat{z}$ and it has a small additional periodic modulation. For such quasi-homogeneous magnetic fields, the Zeeman interaction can no longer induce trapping; instead, it reverts to its traditional role of lifting the spin degeneracy. The system still corresponds to a $\phi/\phi_0 = 1/2$ Hofstadter butterfly, but now in the other asymptotic limit, that is, that of a weak periodic modulation. In this case, each Landau level splits into q sub-bands, with a bandwidth controlled by the amplitude of the weak modulation. The case $\phi/\phi_0 = 1/2$ has $q = 1$, that is, there are no additional gaps in the band structure. We therefore expect to see the usual integer quantum Hall effect in magnetotransport in this regime¹⁷.

We emphasize three aspects of the present work. (1) To our

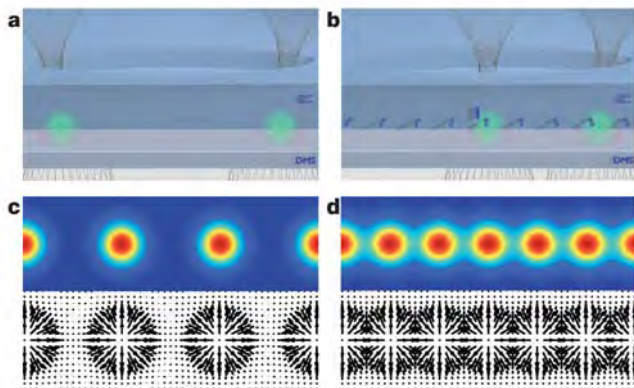


Figure 2 Devices with tunable distance between spin-charge textures. **a**, A finite-size groove is produced in the SC film by etching, and two vortices are trapped in it. At equilibrium, the vortices will stay at maximum possible separation owing to repulsion. **b**, If a current flows through the SC layer, its transversal force on the vortices will move them (and their spin textures) closer together. **c**, Field (top panel) and spin texture (bottom panel) for a 1D vortex array in the low field regime, when inter-vortex separation $a(B)$ is larger than the penetration depth, $a(B) > \lambda$. Vortices trap bound states and form textures independently of each other; there is little overlap between the resulting spin and charge textures. **d**, Same as in **c**, but now for the intermediate-to-high magnetic field regime: there is substantial overlap between the bound states, resulting in a collective spin and charge texture.

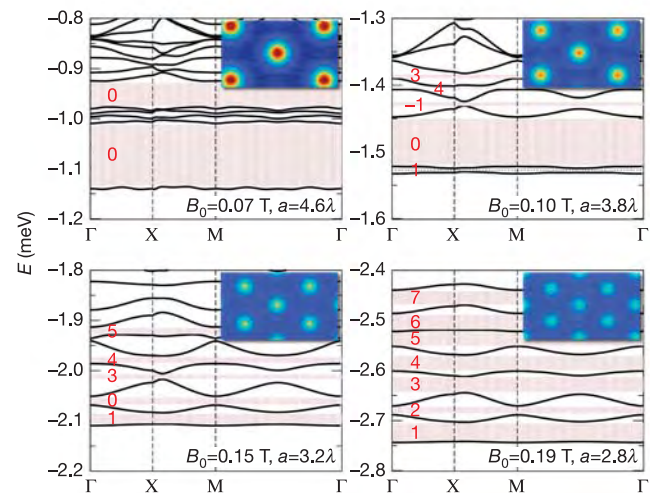


Figure 3 Band structure of the DMS in the hybrid system for an external magnetic field $B_0 = 0.07, 0.10, 0.15$ and 0.19 T. The symmetry points in the Brillouin zone are defined as usual: $\Gamma = (0,0)$, $M = (\pi/a,0)$ and $X = (\pi/a,\pi/a)$. The shaded regions mark the gaps. The value n for the expected plateau in the Hall conductivity, $\sigma_{xy} = ne^2/h$, is marked for the first few gaps. The parameters are as in Fig. 1, except for $z = 8$ nm. The insets show the modulation of the magnetic field in the DMS. The number of vortices per area increases for increasing magnetic fields. a , inter-vortex separation; λ , penetration depth.

knowledge, this is the first proposal of an experimental set-up for the measurement of the Hofstadter butterfly in the limit of a strong periodic modulation (the tight-binding limit). (2) The applied field B plays a very unusual role here. Typically, a fixed lattice cell is assumed. Varying B changes the flux ϕ through the unit cell, resulting in the self-similar eigenspectrum of the Hofstadter butterfly. In contrast, here a change in B implies a change in the lattice constant, but ϕ always equals $\phi_0/2$. As long as there is a vortex lattice, the set-up corresponds to a $\phi/\phi_0 = 1/2$ Hofstadter butterfly irrespective of the value of B . Instead, B controls the amplitude of the periodic modulation, from being the large energy scale (small B) to being a small perturbation (large B). (3) Here the spin plays an essential role, as the periodic modulation is mainly due to the spin-dependent Zeeman potential, which is strongly enhanced in the DMS. In previous work considering periodically modulated magnetic fields^{18–20}, the Zeeman term was ignored. The spin-polarized bound states that we discuss cannot appear in the absence of the Zeeman term, and the physics is qualitatively different.

We now present numerical results that confirm these expectations (for details, see Supplementary Information). In Fig. 3 we show the evolution of the charge carrier energy spectra as B is varied. The large- B limit corresponds to a small periodic modulation, and indeed we see the emergence of equally spaced Landau levels with a weak dispersion due to the weak modulation. As B is lowered, the band structure evolves into that corresponding to a strong modulation, and for the lowest B we see flat bands corresponding to isolated vortex bound levels. This significant change in the dispersion as B is varied is reflected in magnetotransport, in particular in the Hall conductivity. When the Fermi level is inside a gap, the Hall conductivity has a plateau at a quantized value²¹ $\sigma_{xy} = ne^2/h$. Laughlin's gauge arguments²² ensure that this quantization holds in the presence of moderate disorder (small enough to not close these gaps). In the weak-modulation limit, σ_{xy} increases by e^2/h after each transition through a Landau level, as in the typical integer quantum Hall effect. In the tight-binding limit, it is known that the Hall conductance of the entire band is always zero, although if inner gaps are opened by the periodic modulation, they can have different Hall conductances. By explicitly counting the number of zeros (and their

vorticities) of the Bloch wavefunction $u_{\mathbf{k}}(\mathbf{r}) = e^{-ik \cdot \mathbf{r}} \psi_{\mathbf{k}}(\mathbf{r})$ inside the magnetic Brillouin zone^{23–25}, we have verified that the Hall conductance of the sub-bands follows this expected behaviour in both limits of the Hofstadter problem (see Fig. 3).

Figure 4 shows σ_{xy} as a function of the charge carrier density n and magnetic field B . Disorder leads to some widening (arbitrarily drawn in this figure) of each gap, through localization of some fraction of states. The main sources of disorder are inhomogeneity in the DMS and vortex lattice irregularities. The former can be minimized using digital doping, leading to reasonably high mobilities¹⁰. The latter and the so-called Bean profile can be practically eliminated by using SC with low intrinsic pinning. As B is varied at constant n (or vice versa), we expect to see non-trivial sequences of the plateau values of σ_{xy} . Experimental measurement of such non-trivial sequences would clearly signal the emergence of the tight-binding limit of the Hofstadter butterfly, correlated with the appearance of the spin and charge textures in the DMS quantum well.

The present system can be used to investigate many basic physics phenomena, such as Bloch oscillations or the $\phi/\phi_0 = 1/2$ Hofstadter butterfly, but also 1D and 2D metal–insulator transitions, and so on. Recent progress in manipulation of SC vortices can be directly used to create spin lenses, ratchets, pumps, cellular automata, and other devices. Thus, hybrid systems of DMS and SC have remarkable potential for use in both applied and basic condensed matter physics. □

Received 21 July 2004; accepted 10 March 2005; doi:10.1038/nature03559.

1. Peercy, P. S. The drive to miniaturization. *Nature* **406**, 1023–1026 (2000).
2. Wolf, S. A. et al. Spintronics: A spin-based electronics vision for the future. *Science* **294**, 1488–1495 (2001).
3. Lee, C. S., Janko, B., Derenyi, I. & Barabasi, A. L. Reducing vortex density in superconductors using the 'ratchet effect'. *Nature* **400**, 337–340 (1999).
4. Sav'el'ev, S. & Nori, F. Experimentally realizable devices for controlling the motion of magnetic flux quanta in anisotropic superconductors. *Nature Mater.* **1**, 179–184 (2002).
5. Abrikosov, A. A. On the magnetic properties of superconductors of the second group. *Sov. Phys. JETP* **5**, 1174–1182 (1957).
6. Furdyna, J. K. Diluted magnetic semiconductors. *J. Appl. Phys.* **64**, R29–R64 (1988).
7. Berciu, M. & Janko, B. Nanoscale Zeeman localization of charge carriers in diluted magnetic semiconductor-permalloy hybrids. *Phys. Rev. Lett.* **90**, 246804 (2003).
8. Pearl, L. Current distribution in superconducting films carrying quantized fluxoids. *Appl. Phys. Lett.* **5**, 65–66 (1964).
9. Kawakami, R. K. et al. (Ga,Mn)As as a digital ferromagnetic heterostructure. *Appl. Phys. Lett.* **77**, 2379–2381 (2000).
10. Jaroszynski, J. et al. Ising quantum Hall ferromagnet in magnetically doped quantum wells. *Phys. Rev. Lett.* **89**, 266802 (2002).
11. Bloch, F. The quantum mechanics of electrons in crystal lattices. *Z. Phys.* **52**, 555–600 (1928).
12. Hofstadter, D. R. Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields. *Phys. Rev. B* **14**, 2239–2249 (1976).
13. Claro, F. H. & Wannier, G. H. Magnetic subband structure of electrons in hexagonal lattices. *Phys. Rev. B* **19**, 6068–6074 (1979).
14. Melinte, S. et al. Laterally modulated 2D electron system in the extreme quantum limit. *Phys. Rev. Lett.* **92**, 036802 (2004).
15. Schlosser, T. et al. Internal structure of a Landau band induced by a lateral superlattice: A glimpse of Hofstadter's butterfly. *Europhys. Lett.* **33**, 683–688 (1996).
16. Albrecht, C. et al. Evidence of Hofstadter's fractal energy spectrum in the quantized Hall conductance. *Phys. Rev. Lett.* **86**, 147–150 (2001).
17. von Klitzing, K., Dorda, G. & Pepper, M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.* **45**, 494–497 (1980).
18. Bending, S. J., von Klitzing, K. & Ploog, K. Weak localization in a distribution of magnetic-flux tubes. *Phys. Rev. Lett.* **65**, 1060–1063 (1990).
19. Nielsen, M. & Hedegard, P. Two-dimensional electron transport in the presence of magnetic flux vortices. *Phys. Rev. B* **51**, 7679–7699 (1994).
20. Chang, M. C. & Niu, Q. Electron band structure in a two-dimensional periodic magnetic field. *Phys. Rev. B* **50**, 10843–10850 (1994).
21. Thouless, D. J., Kohmoto, M., Nightingale, M. P. & Dennis, M. Quantized Hall conductance in a two-dimensional periodic potential. *Phys. Rev. Lett.* **49**, 405–408 (1982).
22. Laughlin, R. B. Quantized Hall conductivity in 2 dimensions. *Phys. Rev. B* **23**, 5632–5633 (1981).
23. Kohmoto, M. Topological invariant and the quantization of the Hall conductance. *Ann. Phys.* **160**, 343–354 (1985).
24. Usov, N. A. Theory of the quantum Hall effect in a two-dimensional periodic potential. *Sov. Phys. JETP* **67**, 2565–2573 (1988).
25. Demikhovskii, V. Y. & Khomitskiy, D. V. Quantum Hall effect in a p-type heterojunction with a lateral surface superlattice. *Phys. Rev. B* **68**, 165301 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

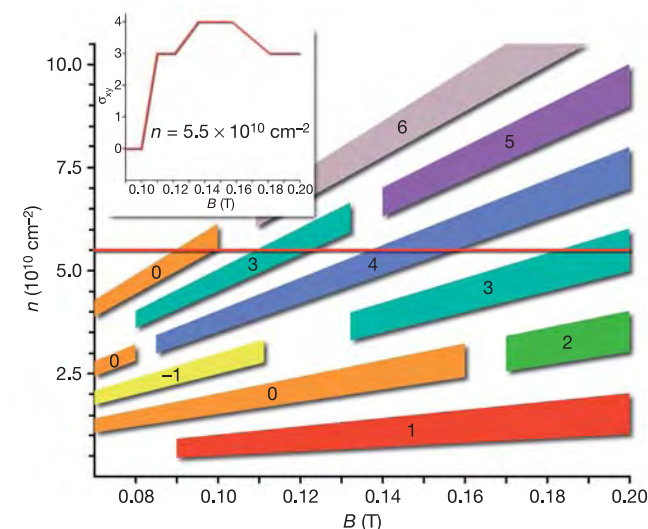


Figure 4 Hall conductivity σ_{xy} , in units of e^2/h , as a function of the magnetic field B and the charge carrier density n . The coloured areas in the main panel mark the gaps between the bands. For some ranges of B , neighbouring bands touch and some of the gaps close. The integers give the quantized values of the σ_{xy} plateaus. For example, σ_{xy} as a function of B , at a constant density $n = 5.5 \times 10^{10} \text{ cm}^{-2}$ (red line) is shown in the inset. It is quantized every time the Fermi level is inside a gap. The parameters are the same as for Fig. 3.

Acknowledgements We thank T. Wojtowicz and A. A. Abrikosov for discussions. This research was supported by NSERC and the Research Corporation (M.B.), by NSF NIRT (T.G.R. and B.J.) and by the Alfred P. Sloan Foundation (B.J.). We are grateful for the hospitality of the Argonne National Laboratory where parts of this work were carried out.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.J. (bjanko@nd.edu).

Topological versus chemical ordering in network glasses at intermediate and extended length scales

Philip S. Salmon¹, Richard A. Martin¹, Philip E. Mason^{2*} & Gabriel J. Cuello³

¹Department of Physics, University of Bath, Bath BA2 7AY, UK

²HH Wills Physics Laboratory, Royal Fort, University of Bristol, Bristol, BS8 1TL, UK

³Institut Laue-Langevin, BP 156, F-38042, Grenoble Cédex 9, France

* Present address: Department of Food Science, Stocking Hall, Cornell University, Ithaca, New York 14853, USA

Atomic ordering in network glasses on length scales longer than nearest-neighbour length scales has long been a source of controversy^{1–6}. Detailed experimental information is therefore necessary to understand both the network properties and the fundamentals of glass formation. Here we address the problem by investigating topological and chemical ordering in structurally disordered AX₂ systems by applying the method of isotopic substitution in neutron diffraction to glassy ZnCl₂. This system may be regarded as a prototypical ionic network forming glass, provided that ion polarization effects are taken into account⁷, and has thus been the focus of much attention^{8–14}. By experiment, we show that both the topological and chemical ordering are described by two length scales at distances greater than nearest-neighbour length scales. One of these is associated with the intermediate range, as manifested by the appearance in the measured diffraction patterns of a first sharp diffraction peak at 1.09(3) Å^{−1}; the other is associated with an extended range, which shows ordering in the glass out to 62(4) Å. We also find that these general features are characteristic of glassy GeSe₂, a prototypical covalently bonded network material^{15,16}. The results therefore offer structural insight into those length scales that determine many important aspects of supercooled liquid and glass phenomenology¹¹.

Binary systems that form network glasses are at the heart of many materials of significant scientific and technological importance¹⁷. A detailed knowledge of the structure is therefore necessary if basic network properties are to be understood—such as their behaviour under extremes of pressure and temperature^{13,18,19}—and materials with new functional properties are to be designed^{20–21}. For glasses, however, a precise description of this structure over a wide range of length scales is notoriously difficult to obtain because the atomic arrangement is not an ordered lattice²². Instead, the atomic sites form a topologically disordered network, and the presence of two chemical species adds further complexity. The identity of the atom occupying a particular site needs to be specified—that is, information is also required on the chemical ordering and hence on how the concentration of a particular species varies across the network (Fig. 1).

Since the pioneering work of Zachariasen²³, it is usual when making models of network glasses to explicitly address the problem of chemical ordering only at a local scale, that is, when identifying structural motifs that act as the basic building blocks such as A(X_{1/2})₄ tetrahedra in AX₂ glasses, where A denotes an electropositive chemical species such as Zn, Ge or Si and X denotes an electronegative chemical species such as Cl, Se or O (refs 2, 24). Workers then focus on the network topology as described by the connectivity of the structural motifs and the features thus generated on intermediate length scales—that is, distances associated with a few nearest neighbours. This approach to modelling glasses is understandable because most experimental probes of the atomic-scale structure are sensitive to the chemical ordering only at relatively short distances. We wish to investigate the extent and influence of chemical ordering on the glass structure at distances associated with a few nearest neighbours. Moreover, does the chemical and topological ordering in homogeneous glasses extend well beyond this intermediate range?

In a neutron-diffraction experiment on an AX₂ system, the measured coherent intensity can be analysed in terms of its contributions from the so-called Bhatia–Thornton²⁵ number–number (N–N), concentration–concentration (C–C) and number–concentration (N–C) partial structure factors, denoted by $S_{NN}^{BT}(k)$, $S_{CC}^{BT}(k)$ and $S_{NC}^{BT}(k)$ respectively, where k is the magnitude of the scattering vector. These structure factors separate the number density from the concentration fluctuations and hence information on the topological order, as described by $S_{NN}^{BT}(k)$, from information on the chemical order, as described by $S_{CC}^{BT}(k)$. The correlation between the number and concentration fluctuations is described by $S_{NC}^{BT}(k)$. The thermodynamic properties of a binary system are readily related to the small- k limit of these functions²⁵: the propensity, for example, of a mixture to form compounds or to phase separate is revealed by the composition dependence of $S_{CC}^{BT}(0)$. For most systems, conventional neutron and X-ray diffraction experiments are particularly sensitive²⁶ to $S_{NN}^{BT}(k)$. However, in a few cases suitable isotopes exist and the full set of partial structure factors can then be separated experimentally by using the powerful method of isotopic substitution in neutron diffraction^{9,15,16}. It is thus feasible to probe both the chemical and topological order over a wide range of length scales.

The measured partial structure factors for glassy ZnCl₂ are plotted in Fig. 2. Collectively, their profiles are similar to those previously measured⁹ for the liquid phase of ZnCl₂ and contrast with those found for a wide range of other ionic AX₂ liquids that do not form glasses²⁶. For example, $S_{NN}^{BT}(k)$ displays three well-defined peaks at the lowest k values, the first at $k_{FSDP} = 1.09(3) \text{ Å}^{-1}$ (the number in parentheses gives the error on the last significant figure: $\pm 0.03 \text{ Å}^{-1}$ in this case), which is called the first sharp diffraction peak (FSDP) and is also observed for $S_{CC}^{BT}(k)$ and $S_{NC}^{BT}(k)$. This feature is associated with intermediate range order^{1–5} having a periodicity $2\pi/k_{FSDP} = 5.8(2) \text{ Å}$ and a coherence length $2\pi/\Delta k_{FSDP} = 12.6(3) \text{ Å}$, where Δk_{FSDP} is the full-width at half-maximum of the FSDP. Evidently, both topological and chemical ordering occurs on an intermediate length scale.

In real space, the glass structure can readily be described in terms of the partial pair distribution functions for the individual atomic species, namely $g_{ZnCl}(r)$, $g_{ZnZn}(r)$ and $g_{ClCl}(r)$, which are plotted in Fig. 3. These functions give a measure of the probability of finding two atoms of type α and β at a distance r apart. A structure based on tetrahedral Zn(Cl_{1/2})₄ building blocks is thus found, in which there are 3.9(1) Zn–Cl nearest-neighbours at 2.28(1) Å and the Cl–Cl distance is 3.70(1) Å. The tetrahedra share corners to give 4.0(1) Zn–Zn nearest neighbours at a distance of 3.75(1) Å and there is an average of 12.1(2) Cl–Cl nearest neighbours over the range $3.19 \leq r(\text{Å}) \leq 4.97$. The packing fraction of the Cl[−] ions (of diameter 3.70 Å) is 0.635, close to the value of 0.64 found for the random packing of hard

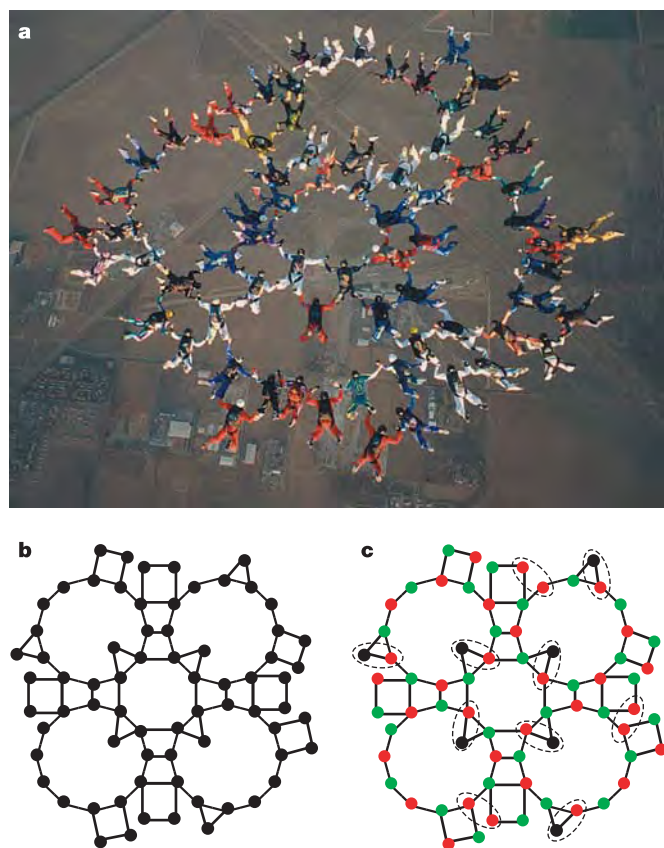


Figure 1 An illustration of topological versus chemical ordering in disordered network systems. **a**, A formation of skydivers helps to demonstrate disorder on an intermediate length scale²¹. Each diver has a simple set of rules for bonding to the next, but there is sufficient flexibility for different patterns of ordering to be created on the scale of a few body lengths. Photo from ref. 21 copyright Mike Skeffington (www.skeff.com). **b**, The arrangement can be made more symmetrical but the sites of the idealized network do not display translational periodicity, that is, it is not possible to take a subset of the diver positions (marked by solid black circles) and translate the resultant coordinates to generate the entire formation. The rotational symmetry of the arrangement is destroyed as divers fasten on at larger distances from the centre. Indeed, it is easy to envisage a more disordered network developing as it expands but the length scale over which complete disorder ensues is less easy to appreciate. In a diffraction experiment the network sites are described by $g_{NN}^{BT}(r)$. **c**, For a binary system of 'red' and 'green' divers, the majority of network sites can be decorated such that unlike colours are nearest neighbours, that is, there is local colour (or chemical) ordering. However, sites can be identified (solid black circles) for which this rule cannot be satisfied. Decorating these remaining sites with red divers then leads to the emergence of a colour-ordering 'superstructure' associated with pairs of red divers on an intermediate length scale. In diffraction experiments this colour (or chemical) ordering is described by $g_{CC}^{BT}(r)$, whereas the correlation between a site and its occupancy by a given colour is described by $g_{NC}^{BT}(r)$.

spheres. The local structure is therefore consistent with that previously deduced by ref. 10.

To identify the features contributing to the topological and chemical ordering at longer length scales it is necessary to interpret the real-space structure in terms of the Bhatia–Thornton partial pair distribution functions $g_{NN}^{BT}(r)$, $g_{CC}^{BT}(r)$ and $g_{NC}^{BT}(r)$, which are shown in the inset to Fig. 3. Here $g_{NN}^{BT}(r)$ describes the sites of the scattering nuclei but does not distinguish between the chemical species that decorate those sites, $g_{CC}^{BT}(r)$ describes the chemical ordering and has positive or negative peaks when there is a preference for like or unlike neighbours respectively, and $g_{NC}^{BT}(r)$ describes the correlation between sites and their occupancy by a given chemical species²⁶. We find for glassy $ZnCl_2$ that $g_{CC}^{BT}(r)$ does

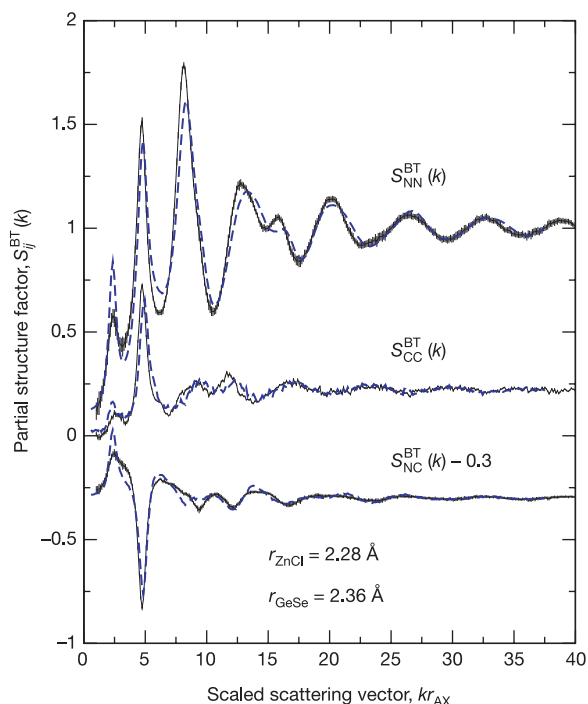


Figure 2 The measured Bhatia–Thornton partial structure factors for glassy $ZnCl_2$ and $GeSe_2$. The $S_{ij}^{BT}(k)$ functions for $ZnCl_2$ (solid curves with vertical bars; the bars give the statistical errors on the data points and are comparable to the curve thickness at most k values) and $GeSe_2$ (broken curves) are plotted as a function of the scaled scattering vector kr_{AX} where r_{AX} is the separation of the unlike nearest neighbours. Each function has an FSDP at $\sim 1 \text{ \AA}^{-1}$ and $S_{NN}^{BT}(k)$ is characterized by three well-defined peaks for $0 \leq kr_{AX} \leq 11$.

not correspond at short and intermediate distances to a simple damped oscillatory function of the type that is found for AX_2 molten salts when there is no FSDP but an intense principal peak at $k_{PP} \approx 2 \text{ \AA}^{-1}$ in $S_{CC}^{BT}(k)$ (ref. 26). Instead, $g_{CC}^{BT}(r)$ has a more complex profile at intermediate distances, which we will refer to as a chemical ordering superstructure. At extended distances of $r > 12 \text{ \AA}$, however, simple damped oscillatory behaviour is recovered that corresponds to alternate shells of like and unlike neighbours. Furthermore, at these extended length scales we also find oscillations in $g_{NN}^{BT}(r)$ and $g_{NC}^{BT}(r)$ that persist to a distance of $62(4) \text{ \AA}$; this far exceeds the coherence length estimated from the width of the FSDP (Fig. 4). In addition, the periodicity of these extended range oscillations is not determined by the position of the FSDP⁴ but by the position of the principal peak at $k_{PP} = 2.10(1) \text{ \AA}^{-1}$ in Fig. 2 which gives $2\pi/k_{PP} = 2.99(2) \text{ \AA}$. The FSDP therefore has little influence on the extended range order: truncating $g_{ij}^{BT}(r)$ at $r \geq 12 \text{ \AA}$ and Fourier transforming has little effect on the observed FSDP in the corresponding $S_{ij}^{BT}(k)$. If a correction is made for the finite k -space resolution function of the neutron diffractometer²⁷, the amplitude of the extended range oscillations is increased.

It is natural to ask whether the general features observed in the chemical and topological ordering are characteristic of other AX_2 network forming glasses in which the bonding scheme differs from that of $ZnCl_2$. The only other system for which sufficiently accurate partial structure factors have been measured is glassy $GeSe_2$, which can be regarded as a prototypical covalently bonded network material^{15,16}. For example, $GeSe_2$ shows typical semiconducting behaviour and the ionicity of the A–X bond on the Pauling electronegativity scale is 7% for $GeSe_2$, as opposed to 43% for $ZnCl_2$. Furthermore, although the predominant structural motif in glassy $GeSe_2$ is also the $A(X_{1/2})_4$ tetrahedron, edge-sharing

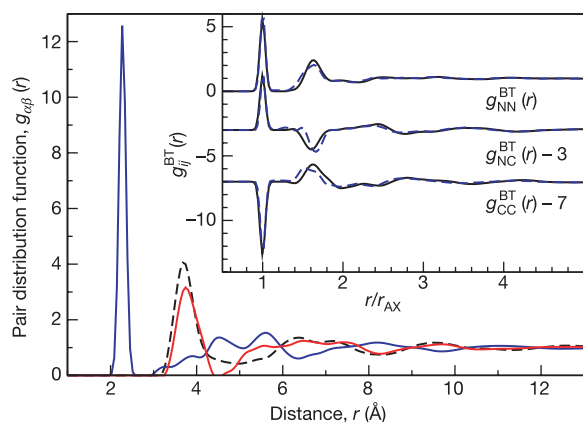


Figure 3 The measured partial pair distribution functions for glassy ZnCl_2 and GeSe_2 . The main panel shows the functions $g_{\text{ZnCl}}(r)$ (solid blue curve), $g_{\text{ZnZn}}(r)$ (solid red curve) and $g_{\text{ClCl}}(r)$ (broken black curve) for ZnCl_2 . The inset shows the Bhatia–Thornton partial pair distribution functions, $g_{ij}^{\text{BT}}(r)$, for ZnCl_2 (solid black curves) and GeSe_2 (broken blue curves) plotted as a function of the scaled distance r/r_{AX} , which ensures that the first peak positions are superimposed. The $g_{ij}^{\text{BT}}(r)$ are obtained from the $g_{\alpha\beta}(r)$ by using the relations $g_{\text{NN}}^{\text{BT}}(r) = c_{\text{A}}^2 g_{\text{AA}}(r) + c_{\text{X}}^2 g_{\text{XX}}(r) + 2c_{\text{AX}} g_{\text{AX}}(r)$, $g_{\text{CC}}^{\text{BT}}(r) = c_{\text{AX}}[g_{\text{AA}}(r) + g_{\text{XX}}(r) - 2g_{\text{AX}}(r)]$ and $g_{\text{NC}}^{\text{BT}}(r) = c_{\text{A}}[g_{\text{AA}}(r) - g_{\text{AX}}(r)] - c_{\text{X}}[g_{\text{XX}}(r) - g_{\text{AX}}(r)]$.

configurations are common and homopolar bonds are a characteristic feature that enables glass formation in the Ge–Se system over a wide composition range. The reciprocal- and real-space data sets for glassy ZnCl_2 and GeSe_2 are compared in Figs 2–4. It is found that the overall features associated with the chemical and topological ordering in both systems are very similar. For example, all three $S_{ij}^{\text{BT}}(k)$ for glassy GeSe_2 show an FSDP at $1.00(2) \text{ \AA}^{-1}$, the chemical ordering superstructure on the intermediate scale is comparable to glassy ZnCl_2 , there is extended range ordering having a periodicity of $\sim 2\pi/k_{\text{PP}}$, and the topological and chemical ordering are correlated on all length scales.

The different bonding schemes for ZnCl_2 and GeSe_2 and the corresponding variation in properties such as their fragility must therefore manifest themselves as subtle but important differences in structural detail. In the Angell classification scheme^{11,18}, a taxonomy which stems from the temperature dependence of the liquid viscosity, ZnCl_2 is regarded as intermediate between ‘strong’ glass formers such as SiO_2 and ‘fragile’ glass formers such as the molten mixture $3\text{KNO}_3\text{--}2\text{Ca}(\text{NO}_3)_2$. By comparison, GeSe_2 is a fragile liquid at high temperatures but becomes increasingly strong as the glass-transition temperature is approached²⁸. The present results show that the glassy phase has two characteristic length scales at distances larger than the nearest neighbour. One is associated with the intermediate range, as manifested by the appearance of an FSDP, and the other is associated with an extended range, which has a strong contribution from the principal peak and thus relates to a propagation of short range ordering.

It is therefore plausible that, as a liquid is supercooled, signatures of its fragility are frozen into the glass structure and are represented by features such as the sharpness of the FSDP and spatial persistence of the extended range ordering. The fragility of systems such as ZnCl_2 and GeSe_2 will then be sensitive to the rate at which the structure of the supercooled liquid on these two length scales changes with temperature. Also, the comparative importance of the intermediate and extended range ordering may give structural insight into other important phenomena such as the nucleation and growth of crystals from the glassy or liquid phase^{11,13,14} and the existence or otherwise of polyamorphic phase transitions, that is, distinct changes in the structure of a network liquid or glass with change of density^{13,18,19}. Glasses have thus begun to reveal the secrets

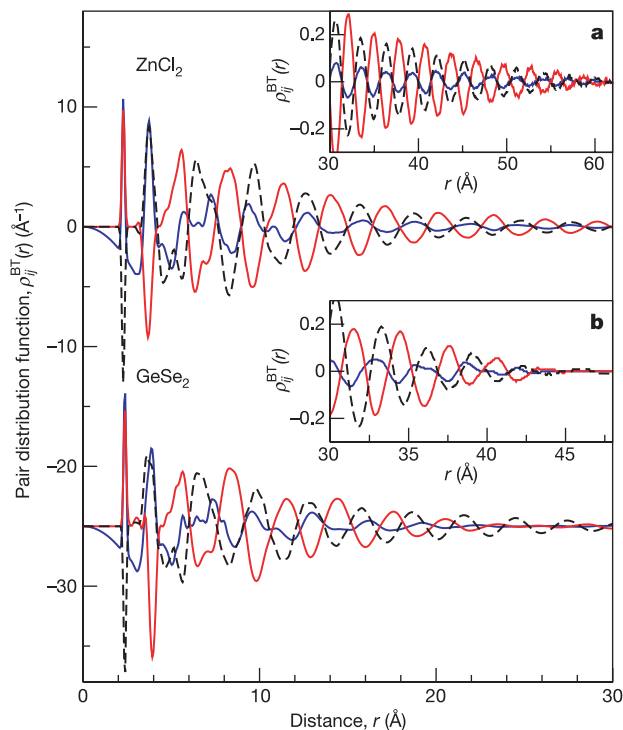


Figure 4 Extended range ordering is found for glassy ZnCl_2 and GeSe_2 . This ordering is emphasized by plotting the Bhatia–Thornton partial pair distribution functions for ZnCl_2 (upper set of curves) and GeSe_2 (lower set of curves shifted down by $25 \text{ \AA}^{-1}$) in the form $\rho_{\text{NN}}^{\text{BT}}(r) = 4\pi n_0 r^2 [g_{\text{NN}}^{\text{BT}}(r) - 1]$ (solid blue curves), $\rho_{\text{CC}}(r) = 4\pi n_0 r^2 g_{\text{CC}}^{\text{BT}}(r)$ (broken black curves) and $\rho_{\text{NC}}(r) = 4\pi n_0 r^2 g_{\text{NC}}^{\text{BT}}(r)$ (solid red curves) where n_0 is the atomic number density. The insets show the data for ZnCl_2 (a) and GeSe_2 (b) at large distances on an expanded scale.

of their structure on length scales that are relevant for understanding several outstanding issues in glass science. □

Methods

The synthesis of ZnCl_2

Isotopic H^*Cl was prepared by the action of H_2SO_4 on the relevant isotopic CsCl in a high vacuum apparatus. The HCl was then reacted with an aqueous solution of AgNO_3 (99.995%) to precipitate AgCl , which was washed in water and dried. Zinc metal (99.999%, 0.45 mol equivalent) was distilled under high vacuum into a sealed silica ampoule that contained the AgCl (1 mol equivalent). After 24 h at 700°C the zinc had completely reacted with the AgCl and the resultant ZnCl_2 was condensed in a cylindrical silica ampoule (inside diameter 4 mm, wall thickness 1 mm). The ampoule containing liquid ZnCl_2 was quenched from 400°C in water at room temperature to give a clear colourless glass. The ampoule was then sealed under high vacuum ready for use in the diffraction experiments. An AgCl impurity content of $<0.1 \text{ mol\%}$ was estimated by chemical analysis and energy-dispersive neutron-scattering experiments showed no sign of the strong ^{109}Ag absorption resonance at $5.19(1) \text{ eV}$. A neutron-diffraction experiment on a crystallized sample revealed the anhydrous delta phase¹⁰ of ZnCl_2 .

Partial structure factors

The diffraction experiments were made at $25(1)^\circ\text{C}$ on samples of glassy $\text{Zn}^{35}\text{Cl}_2$, $\text{Zn}^{\text{nat}}\text{Cl}_2$, $\text{Zn}^{\text{mod}}\text{Cl}_2$ and $\text{Zn}^{37}\text{Cl}_2$ contained in silica ampoules using the D4C instrument (Institut Laue-Langevin) with an incident neutron wavelength of $0.49971(1) \text{ \AA}$. The ^{35}Cl : ^{37}Cl ratio for each of the samples was $0.994:0.006$, $0.7577:0.2423$, $0.42:0.58$ and $0.018:0.982$, respectively. Diffraction patterns were taken for the samples in their container, the empty container, and a cylindrical vanadium rod for normalization purposes. The intensity for a bar of neutron absorbing $^{10}\text{B}_4\text{C}$ of dimensions comparable to the sample was also measured to account for the effect of sample self-shielding on the background count rate at small scattering angles. The data were carefully corrected to yield the total structure factor for each sample and the usual self-consistency checks were performed¹⁶. The total structure factor obtained for the glassy $\text{Zn}^{\text{nat}}\text{Cl}_2$ sample was in excellent agreement with that previously measured by ref. 10 (see Supplementary Figs S1 and S2). The partial structure factors were calculated from the total structure factors by using the method of singular value decomposition²⁹. The neutron-scattering lengths were taken from ref. 30 and the atomic number density of the glass⁸ is $0.0359(1) \text{ \AA}^{-3}$. The $g_{\alpha\beta}(r)$ were obtained by spline-fitting and Fourier-transforming the corresponding partial structure factors,

$S_{\alpha\beta}(k)$, after truncating either (1) abruptly at $k_{\max} = 23.5 \text{ \AA}^{-1}$ by using a step function $M(k) = 1$ for $k \leq k_{\max}$, $M(k) = 0$ for $k > k_{\max}$ or (2) smoothly by using, for example, a Lorch function $M(k) = \sin(\pi k/k_{\max})/(\pi k/k_{\max})$. Procedure (1) gives well-defined peaks in r -space but in the case of $g_{\text{ZnCl}}(r)$ requires a correction for the effect of $M(k)$ (refs 15, 16). Procedure (2) gives smoother $g_{\alpha\beta}(r)$ functions at all r values but leads to a loss in resolution of the first peaks. The $g_{\alpha\beta}(r)$ function obtained by using procedure (1) was truncated after the first few peaks and smoothly joined with the corresponding $g_{\alpha\beta}(r)$ obtained by using procedure (2) when the loss of resolution caused by the smoother $M(k)$ could no longer be discerned. The unphysical low- r oscillations were then set to zero and the final $g_{\alpha\beta}(r)$ was Fourier-transformed into k -space to confirm agreement with the original $S_{\alpha\beta}(k)$ (see Supplementary Fig. S3).

Received 19 January; accepted 16 February 2005; doi:10.1038/nature03475.

1. Moss, S. C. & Price, D. L. in *Physics of Disordered Materials* (eds Adler, D., Fritzsche, H. & Ovshinsky, S. R.) 77–95 (Plenum, New York, 1985).
2. Elliott, S. R. Medium-range structural order in covalent amorphous solids. *Nature* **354**, 445–452 (1991).
3. Salmon, P. S. Real space manifestation of the first sharp diffraction peak in the structure factor of liquid and glassy materials. *Proc. R. Soc. Lond. A* **445**, 351–365 (1994).
4. Elliott, S. R. Extended-range order, interstitial voids and the first sharp diffraction peak of network glasses. *J. Non-Cryst. Solids* **182**, 40–48 (1995).
5. Gaskell, P. H. & Wallis, D. J. Medium-range order in silica, the canonical network glass. *Phys. Rev. Lett.* **76**, 66–69 (1996).
6. Massobrio, C., Celino, M. & Pasquarello, A. Charge fluctuations and concentration fluctuations at intermediate-range distances in the disordered network-forming materials SiO_2 , SiSe_2 , and GeSe_2 . *Phys. Rev. B* **70**, 174202 (2004).
7. Wilson, M. & Madden, P. A. Short- and intermediate-range order in MCl_2 melts: the importance of anionic polarization. *J. Phys. Condens. Matter* **5**, 6833–6844 (1993).
8. Angell, C. A. & Wong, J. Structure and glass transition thermodynamics of liquid zinc chloride from far-infrared, Raman and probe ion electronic and vibrational spectra. *J. Chem. Phys.* **53**, 2053–2066 (1970).
9. Biggin, S. & Enderby, J. E. The structure of molten zinc chloride. *J. Phys. C* **14**, 3129–3136 (1981).
10. Desai, J. A. E., Wright, A. C., Wong, J. & Sinclair, R. N. A neutron diffraction investigation of the structure of vitreous zinc chloride. *J. Non-Cryst. Solids* **51**, 57–86 (1982).
11. Angell, C. A. Spectroscopy simulation and scattering, and the medium range order problem in glass. *J. Non-Cryst. Solids* **73**, 1–17 (1985).
12. Wilson, M. & Madden, P. A. Voids, layers and the first sharp diffraction peak in ZnCl_2 . *Phys. Rev. Lett.* **80**, 532–535 (1998).
13. Polysky, C. H. *et al.* Pressure-induced crystallization of vitreous ZnCl_2 . *Phys. Rev. B* **61**, 5934–5938 (2000).
14. Wilson, M. Nucleation and growth of polypoly-layered crystals from the network liquid zinc chloride. *J. Chem. Phys.* **118**, 9838–9853 (2003).
15. Petri, I., Salmon, P. S. & Fischer, H. E. Defects in a disordered world: the structure of glassy GeSe_2 . *Phys. Rev. Lett.* **84**, 2413–2416 (2000).
16. Salmon, P. S. & Petri, I. Structure of glassy and liquid GeSe_2 . *J. Phys. Condens. Matter* **15**, S1509–S1528 (2003).
17. Boolchand, P. (ed.) *Insulating and Semiconducting Glasses* (World Scientific, Singapore, 2000).
18. Angell, C. A. Formation of glasses from liquids and biopolymers. *Science* **267**, 1924–1935 (1995).
19. Crichton, W. A., Mezouar, M., Grande, T., Stølen, S. & Grzechnik, A. Breakdown of intermediate-range order in liquid GeSe_2 at high pressure. *Nature* **414**, 622–625 (2001).
20. Martin, J. D., Goettler, S. J., Fossé, N. & Iton, L. Designing intermediate-range order in amorphous materials. *Nature* **419**, 381–384 (2002).
21. Salmon, P. S. Order within disorder. *Nature Mater.* **1**, 87–88 (2002).
22. Ziman, J. M. *Models of Disorder* (CUP, Cambridge, 1979).
23. Zachariasen, W. H. The atomic arrangement in glass. *J. Am. Chem. Soc.* **54**, 3841–3851 (1932).
24. Wright, A. C. *et al.* The structure of some simple amorphous network solids revisited. *J. Non-Cryst. Solids* **129**, 213–232 (1991).
25. Bhatia, A. B. & Thornton, D. E. Structural aspects of the electrical resistivity of binary alloys. *Phys. Rev. B* **2**, 3004–3012 (1970).
26. Salmon, P. S. The structure of molten and glassy 2:1 binary systems: an approach using the Bhatia-Thornton formalism. *Proc. R. Soc. Lond. A* **437**, 591–606 (1992).
27. Grimley, D. I., Wright, A. C. & Sinclair, R. N. Neutron scattering from vitreous silica. *J. Non-Cryst. Solids* **119**, 49–64 (1990).
28. Stølen, S., Grande, T. & Johnsen, H.-B. Fragility transition in GeSe_2 -Se liquids. *Phys. Chem. Chem. Phys.* **4**, 3396–3399 (2002).
29. Ludwig, K. F. Jr, Warburton, W. K., Wilson, L. & Bienenstock, A. I. Liquid GeBr_4 . 1. A test of the anomalous x-ray scattering method as applied to disordered materials. *J. Chem. Phys.* **87**, 604–612 (1987).
30. Sears, V. F. Neutron scattering lengths and cross sections. *Neutron News* **3**, 26–37 (1992).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank M. Skeffington for permission to use his photo in Fig. 1a, R. L. McGreevy for assistance with the isotopes, P. Palleau for help with the diffraction experiment, A. C. Barnes for discussions on the SVD method, A. C. Hannon for a powder diffraction pattern on the polycrystalline phase of ZnCl_2 , A. C. Wright for providing his data for glassy ZnCl_2 , P. Buchanan and A. K. Soper for help with an earlier experiment at ISIS, R. Evans, P. A. Madden, C. Massobrio and M. Wilson for discussions, and the EPSRC and Alliance Programme for financial support. P.E.M. made the samples, G.J.C. prepared the diffractometer and R.A.M. made the data analysis. P.S.S. conceived the experiment and was involved in all aspects of the work.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.S.S. (p.s.salmon@bath.ac.uk).

Measurement of femtometre-scale atomic displacements by X-ray absorption spectroscopy

Robert F. Pettifer¹, Olivier Mathon², Sakura Pascarelli², Michael D. Cooke^{3*} & Michael R. J. Gibbs³

¹Department of Physics, University of Warwick, Coventry CV4 7AL, UK

²European Synchrotron Radiation Facility (ESRF), BP 220, 38043 Grenoble Cedex 9, France

³Department of Engineering Materials, Sir Robert Hadfield Building, Mappin Street, University of Sheffield, Sheffield S1 3JD, UK

* Present address: School of Engineering, University of Durham, Durham DH1 3LE, UK

The frequencies of extended X-ray absorption fine-structure (EXAFS)¹ measurements, which are oscillations occurring on the high-energy side of an X-ray absorption edge, can be used to identify interatomic distances in materials. We have used a dispersive X-ray spectrometer^{2–5}, which has no moving components, to make rapid measurements with minimal energy drift of the difference in EXAFS from the Fe K edge in an iron-cobalt thin film undergoing periodic strain through magnetostriction^{6,7}. We show that magnetostriction can be detected by differential X-ray absorption. The magnitude of the recorded signal relative to the noise shows a sensitivity to mean differential atomic motion of one femtometre: a factor of 100 times more sensitive than that normally available^{8,9}.

EXAFS is caused by an internal photoelectron scattering process, which modulates the X-ray absorption coefficient. The present consensus is that the accuracy of interatomic distance determination by EXAFS is between 0.01 Å and 0.001 Å, depending on the circumstances^{8,9}. However, in certain types of study, such as the differential magnetostriction measurements reported here, direct comparison should be capable of much higher precision. This is especially true when synchrotron radiation is used from a third-generation source, where the fluxes available can be 10^{13} photons $\text{s}^{-1} \text{eV}^{-1}$. These should be able to produce relative statistical errors in the absorption spectrum of $\sim 10^{-6}$ under optimal conditions in a few hours. With such a statistical accuracy, interatomic strains of the order of femtometres and below should be detectable with EXAFS.

It is not, however, only the statistical precision that limits the accuracy of interatomic distance determination, it is also the stability of the X-ray energy. Model calculations show that an energy shift of 0.01 eV between two measurements to be compared can give rise to signals comparable to those anticipated from magnetostriction, compromising any comparison at the level reported here. To overcome this problem, we used the dispersive X-ray absorption spectrometer (ID24)^{2–5} of the European Synchrotron Radiation Facility (ESRF). This spectrometer has no moving parts and is thus inherently more stable than a conventional step-scanning instrument, and it allows comparative measurements to be taken rapidly. Our measurements were performed on a thin film of the alloy FeCo located between the poles of a magnet, which produced a saturating field in the sample, Fig. 1 (see Supplementary Information S1). The magnets were rotated via a stepping motor such that measurements of the absorption coefficient were made with the magnetic field in the sample plane, either parallel or perpendicular to the X-ray polarization vector. This arrangement exploits the anisotropy of the photoelectron emission^{10–12}, which lies predominantly along the X-ray polarization vector. As the magnet is rotated around the sample, the induced magnetization, which produces the strain, lies either along or perpendicular to the X-ray polarization vector. Because structural changes are reflected

in the absorption, X-ray linear dichroism results.

Transmitted intensity measurements were made repeatedly at every 90° angle between the magnetization vector and the photon polarization vector. The differential absorption was calculated from the transmitted intensities¹³. The differential fine-structure function, $\Delta\chi_{\pm}$, could then be calculated by dividing by the edge jump (see Supplementary Information S2). An entire four-quadrant measurement took about 1 s, with repeated measurements accumulated over a 2-h period. The experiment was then repeated but this time the magnet was rotated 90° in phase with respect to the previous measurements. This should invert the signal. The two curves, one multiplied by -1 , are shown in Fig. 2. The maximum derivative of the absorption data occurs at the first inflection point of the edge, at 0 eV. No differential signal is seen at this energy, testifying to the beam energy stability. The structure we observe exists well into the continuum state energy regime of the photoelectron, forcing us to conclude that the origin of $\Delta\chi_{\pm}$ is structural.

The X-ray absorption fine-structure function is defined by $\chi(k) = \frac{\mu(k) - \mu_0(k)}{\mu_0(k)}$, where k is the magnitude of the photoelectron wave vector and $\mu(k)$ and $\mu_0(k)$ are the absorption coefficients with and without scattering atoms present for, in this case, K-shell absorption. This function can be written in the simple form^{10,14,15}

$$\chi(k) = \sum_j A_j(k) \sin(kS_j + \Phi_j(k)) \quad (1)$$

where the sum takes place over a number of shells. $A_j(k)$ is an amplitude for the scattering from the j th shell, S_j is the path length starting and ending on the photoabsorbing atom, and $\Phi_j(k)$ is a phase shift associated with the emergence of the photoelectron from the photoabsorber and from the scattering of the atoms in the path. The paths are not restricted to single scattering paths and may include several scattering atoms. The details of the quantum mechanical scattering of the photoelectrons by the atoms in the material are included in the amplitude $A_j(k)$, which also includes thermal effects, and phase $\Phi_j(k)$. Performing a first-order Taylor expansion on equation (1) gives

$$\Delta\chi = \sum_j A_j(k) \cos(kS_j + \Phi_j(k)) k \Delta S_j \quad (2)$$

(see Supplementary Information S3)^{14,16}. Equation (2) shows that the contributions to the amplitudes of the differential signal are scaled differently relative to the normal EXAFS signal, equation (1). First, the differential signal has an additional k weighting, which will

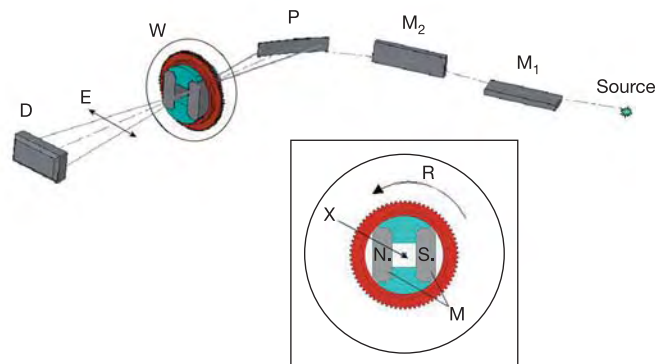


Figure 1 A schematic of the experimental layout. The source is a planar undulator. The X-ray beam from the undulator is focused by mirrors, vertically by M_1 and horizontally by M_2 . The horizontal divergent beam is incident on a bent silicon polychromator P. This focuses the beam onto the sample X, which is held in the centre of a wheel W. Inset, wheel supports permanent magnets, M (N, north; S, south), and these are rotated (R) to position the magnetization in the plane of the sample with respect to the electric field $\mathbf{E}$ of the X-rays. The spectrum is recorded by the 16-bit CCD camera D (ref. 5), which views a phosphor.

enhance the high- k (energy) part of the spectrum, and second, the amplitude of the signal will be proportional to the displacement of each shell. Thus the normal situation for EXAFS of the near neighbours dominating the EXAFS signal is not necessarily true for the differential signal.

In the study of magnetostriction, we exploit the anisotropy of the X-ray absorption of the iron K edge in FeCo to monitor the strain produced by the magnetization using the assumption that the macroscopic magnetostriction tensor holds at the microscopic level. In our particular case the magnetization always results in a positive strain in the direction of the magnetization. We can accommodate the differential anisotropy by modifying equation (2) such that $\Delta\chi$ becomes $\Delta\chi_{\pm}$, the differential EXAFS signal between the magnetization being parallel and perpendicular to the electric field vector of the linearly polarized radiation, and ΔS_j becomes $\langle \Delta S_j^{\pm} \rangle$, the average change of path length for the two polarization cases for the j th shell scattering weighted by the EXAFS polarization dependence. For the case of body-centred-cubic type materials, such as FeCo (ref. 17), the strain in a magnetically saturated material is given by a fourth-rank cartesian tensor that contains three independent coefficients. In a commonly accepted notation⁶, these are, for the ordered bulk FeCo alloy, $\lambda^{(g,2)} = 210$ p.p.m. and $\lambda^{(e,2)} = 60$ p.p.m. (The other coefficient,

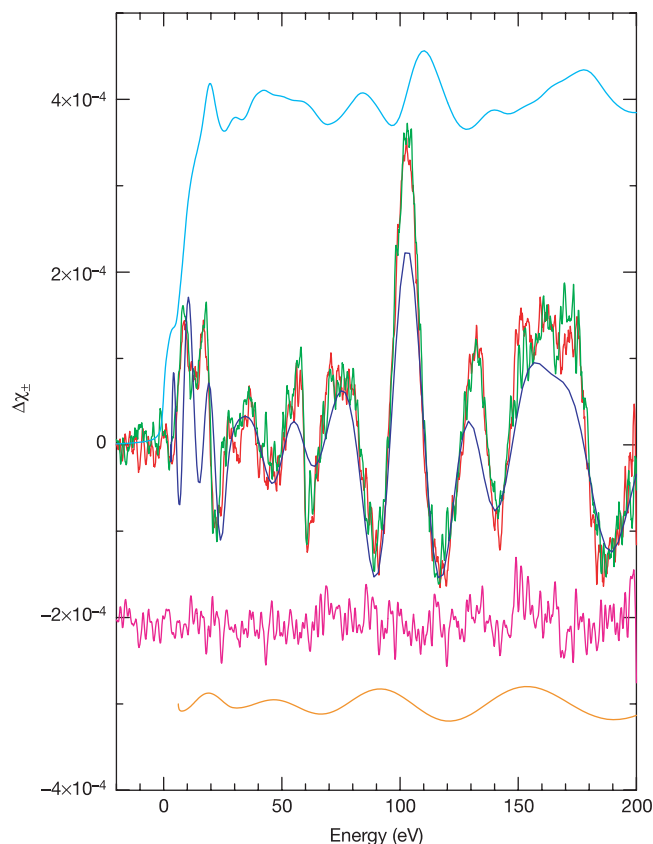


Figure 2 The magnetostrictive differential EXAFS signal. The red curve shows the differential EXAFS created from the absorption with the magnetization $\mathbf{M}$ perpendicular to the electric polarization vector $\mathbf{p}$ subtracted from the absorption with $\mathbf{M}$ parallel to $\mathbf{p}$. In a separate experiment, the data acquisition was started with a 90° phase shift and then inverted to give the green curve. The theoretical calculation (dark blue curve) is shown for comparison. The normalized X-ray absorption of the sample (light blue curve) is also shown; the mean absorption above the edge is set to unity. The energy axis is relative to the first inflection point of the Fe K edge (7,112 eV). The purple line is the difference between the two differential EXAFS measurements divided by $\sqrt{2}$ to represent the noise present in each spectrum. The brown oscillatory line is a theoretical signal from a mean 1 fm displacement of the first shell.

Table 1 Some dominant contributing shells of atoms to the differential fine structure

Shell j	Path length, S_j (Å)	$\langle \Delta S_j^{\pm} \rangle$ (fm)	Number of legs	Scattering from atom*
a	4.97	4.1	2	Co atom at (1/2, 1/2, 1/2)
b	5.73	15.1	2	Fe atom at (1, 0, 0)
c	8.12	10.4	2	Fe atom at (1, 1, 0)
d	9.51	16.9	2	Co atom at (3/2, 1/2, 1/2)
e	11.47	30.1	3	Fe atom at (2, 0, 0) via single scatter from Fe atom at (1, 0, 0)
f	11.47	30.1	4	Fe atom at (2, 0, 0) via double scatter from Fe atom at (1, 0, 0)

The shells are numbered in increasing distance from the absorber. The third column gives the average path length contribution to the differential EXAFS between the magnetization vector being parallel and perpendicular to the photon polarization vector. The only multiple scattering paths that produce a major contribution are those that are collinear, such as paths e and f.

*Relative to Fe atom at the origin in terms of the unit cell of the cubic lattice.

$\lambda^{(\alpha,0)}$, represents a field-dependent volume change and thus does not contribute to the saturation Joule magnetostriction.)

The films on which we make our measurements have a strong texture that points the (110) direction normal to the surface of the film, but the crystallites have a random orientation ϕ around this vector. The averaging to be performed consists of averaging over the sum of all legs of the paths weighted by the polarization sensitivity (see Supplementary Information S4). The dominant paths of the differential fine structure are shown in Table 1.

From equation (2), and using the FEFF8.0 code¹⁸ to calculate $A_j(k)$ and $\Phi_j(k)$, we calculated the differential fine structure seen in Fig. 2 (dark blue curve). It is clear that the main features of the measurement have been reproduced even down to the difficult region below 30 eV, where we know that our theory is inadequate. In addition, the large measured peak at just over 100 eV is reproduced, although not with the same magnitude. The success of the calculation testifies to the structural origin of the signal. Figure 2 also shows the comparison of the noise of the data with a calculated signal that would come from an average interatomic strain in the first shell of 1 fm. If just the first shell structure were visible, then this displacement would be measurable. We can therefore claim that we have sensitivity to mean interatomic displacements of about 1 fm.

The contributions of differing shells to the fine structure are displayed in Fig. 3. The shell-designating letters are the same as those in Table 1 and SS is the sum of contributions from all other shells contributing to the calculation. Surprisingly, the second-shell contribution is larger in amplitude than the first shell. This results from the much larger $\langle \Delta S_2^{\pm} \rangle$ compared with $\langle \Delta S_1^{\pm} \rangle$ (see Table 1), and is a consequence of the relative magnitudes of the saturation magnetostriction tensor elements controlling the strain in the (100) for $\langle \Delta S_2^{\pm} \rangle$ and (111) for $\langle \Delta S_1^{\pm} \rangle$ directions, that is $\lambda^{(g,2)}$ and $\lambda^{(e,2)}$ respectively.

We have assumed that the interatomic strains scale directly with the macroscopic strain. In general, this will not be the case. However, with the realization that EXAFS can provide differential interatomic distance measurements with a resolution of ~ 1 fm, coupled with the power of EXAFS to examine the changes in the local structure surrounding a particular atomic type, new areas of value are opened up. For example, the local magnetostrictive strain surrounding a rare earth element such as terbium or dysprosium in TERFENOL-D, a commercially available giant magnetostrictive material, could verify the local origin of the strain created by the asymmetry of the rare earth atom. Such information is difficult to gather via other techniques. X-ray diffraction, for example, can provide very high lattice parameter accuracy, but it is more difficult to obtain comparable accuracy on differential values of atomic positions within the unit cell. Similarly, X-ray scattering would struggle to provide this type of information from amorphous materials. Further examples of systems that could benefit from

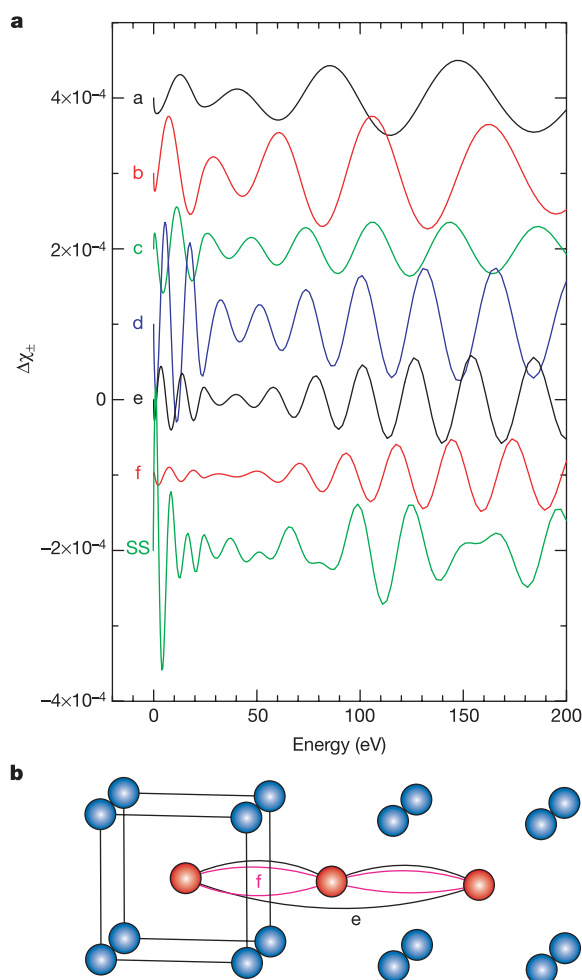


Figure 3 The main contributions to the calculated fine structure. **a**, The individual contributions to the differential fine structure, with the letters referring to the same shells as detailed in Table 1. SS is the sum of contributions other than those above. The curves have been displaced vertically from each other by 1×10^{-4} for clarity. Note that the signal from the second shell is larger than the first. Also note that shells e and f have the same path length, but path e is a double-scattering path whereas path f is a triple-scattering path. The two contributions are almost 180° out of phase. **b**, Diagram illustrating the double (black) and triple (red) scattering paths e and f respectively of the FeCo lattice. Fe atoms are shown brown and Co atoms are blue.

this advance include piezoelectric materials with complex unit cells, elasticity of materials and phase transitions. □

Received 29 October 2004; accepted 4 March 2005; doi:10.1038/nature03516.

- Sayers, D. E., Stern, E. A. & Lytle, F. W. New technique for investigating noncrystalline structures: Fourier analysis of the extended X-ray absorption fine structure. *Phys. Rev. Lett.* **27**, 1204–1207 (1971).
- Hagelstein, M., Fontaine, A. & Goulon, J. High brilliance beamline for time-resolved X-ray absorption spectroscopy at ESRF. *Jpn. J. Appl. Phys.* **32** (suppl. 32–2), 240–242 (1993).
- Hagelstein, M., San Miguel, A., Ressler, T., Fontaine, A. & Goulon, J. The beamline ID24 at the ESRF for energy dispersive X-ray absorption spectroscopy. *J. Phys. IV* **1** (C2), 303–308 (1997).
- Pascarella, S., Mathon, O. & Aquilanti, G. New opportunities for high-pressure X-ray absorption spectroscopy using dispersive optics. *J. Alloys Compounds* **362**, 33–40 (2002).
- Koch, A., Hagelstein, M., San Miguel, A., Fontaine, A. & Ressler, T. in *IS&T/SPIE Proc. Symp. Electronic Imaging: Science and Technology* Vol. 2416 (eds Anagnostopoulos, C. N. & Lesser, M. P.) 85–93 (San Jose, California, 1996).
- du Trémolet de Lacheisserie, E. *Magnetostriction: Theory and Applications of Magnetoelasticity* 133 (CRC, Florida, 1993).
- Gibbs, M. R. J. (ed.) *Modern Trends in Magnetostriction Study and Application* (Kluwer, Dordrecht, 2001).
- Aksenov, V. L., Kuzmin, A. Yu., Purans, J. & Tyutyunnikov, S. I. EXAFS spectroscopy at synchrotron-radiation beams. *Phys. Particles Nuclei* **32**, 1–33 (2001).
- Dalba, G., Fornasini, P., Grisenti, R. & Purans, J. Sensitivity of extended X-ray-absorption fine structure to thermal expansion. *Phys. Rev. Lett.* **82**, 4240–4243 (1999).

10. Lee, P. A. & Pendry, J. B. Theory of the extended X-ray absorption fine structure. *Phys. Rev. B* **11**, 2795–2811 (1975).
11. Brouder, C., Ruiz-Lopez, M. F., Pettifer, R. F., Benfatto, M. & Natoli, C. R. Systematic approach to the calculation of the polarisation dependent (and polarisation averaged) general term of the curved-wave multiple-scattering series in the x-ray absorption cross-section. *Phys. Rev. B* **39**, 1488–1500 (1989).
12. Benfatto, M., Natoli, C. R., Brouder, C., Pettifer, R. F. & Ruiz Lopez, M. F. Polarized curved-wave extended x-ray-absorption fine structure: Theory and application. *Phys. Rev. B* **39**, 1936–1939 (1989).
13. Mathon, O. *et al.* XMCD under pressure at the Fe K edge on the energy dispersive beamline of the ESRF. *J. Synchrotron Rad.* **11**, 423–427 (2004).
14. Benfatto, M. *et al.* Multiple scattering regime and higher-order correlations in x-ray-absorption spectra of liquid solutions. *Phys. Rev. B* **34**, 5774–5781 (1986).
15. Rehr, J. J. & Albers, R. C. Theoretical approaches to x-ray absorption fine structure. *Rev. Mod. Phys.* **72**, 621–654 (2000).
16. Hanham, M. L. & Pettifer, R. F. Generalized Ramsauer-Townsend effect in extended x-ray-absorption fine structure. *Phys. Rev. B* **64**, 180101 (2001).
17. Bayliss, P. Revised unit cell dimensions, space group, and chemical formula of some metallic minerals. *Can. Mineral.* **28**, 751–755 (1990).
18. Ankudinov, A. L., Ravel, B., Rehr, J. J. & Conradson, S. D. Real space multiple scattering calculations of XANES. *Phys. Rev. B* **58**, 7565–7576 (1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We wish to acknowledge R. Weigel, S. Pasternak, A. Sheffield, A. Lovejoy, P. Pinel, M. Borowski and M.-C. Dominguez for their assistance in performing this experiment, or one of its precursors. R.F.P. thanks the ESRF for hosting a sabbatical leave from the University of Warwick and T. F. Jeranko, M. Ruffoni and R. Dupree for discussions.

Authors' contributions R.E.P. conceived the idea, took part in all of the experiments, constructed the theory, performed the calculations and wrote the paper. O.M. took part in all of the experiments, including the unsuccessful precursors and operated ID24. S.P. stabilized ID24, took part in the final measurement and performed major editing of the paper. M.D.C. made the specimens and took part in some of the measurements. M.R.J.G. provided facilities for making the samples, provided magnetostriction expertise and took part in the final measurement.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.E.P. (r.pettifer@warwick.ac.uk).

Slip rate variations on normal faults during glacial–interglacial changes in surface loads

Ralf Hetzel^{1*} & Andrea Hampel²

¹Institute of Geology, Department of Earth Sciences, ETH Zurich, Sonneggstr. 5, 8092 Zurich, Switzerland

²Institute of Geological Sciences, University of Bern, Baltzerstr. 1–3, 3012 Bern, Switzerland

* Present address: Geologisch-Paläontologisches Institut, Universität Münster, Corrensstr. 24, 48149 Münster, Germany

Geologic and palaeoseismological data^{1,2} document a marked increase in the slip rates of the Wasatch fault and three adjacent normal faults in the Basin and Range Province during the Late Pleistocene/Early Holocene epochs³. The cause of this synchronous acceleration of fault slip and the subsequent clustering of earthquakes during the Holocene³ has remained enigmatic, although it has been suggested that the coincidence between the acceleration of slip and the shrinkage of Lake Bonneville after the Last Glacial Maximum may indicate a causal relationship⁴. Here we use finite-element models of a discrete normal fault within a rheologically layered lithosphere to evaluate the relative importance of two competing processes that affect fault slip: postglacial unloading (the removal of mass), which decreases the slip rate, and lithospheric rebound, which promotes faster slip.

We show that lithospheric rebound caused by regression of Lake Bonneville^{4–6} and deglaciation of adjacent mountain ranges^{7,8} provides a feasible mechanism for the high Holocene rates of faulting in the Wasatch region. Our analysis implies that climate-controlled changes in loads applied to Earth's surface may exert a fundamental control on the slip history of individual normal faults.

Many major faults exhibit a more or less cyclic release of elastic strain energy during earthquakes, which led to the proposal of a range of earthquake recurrence models^{9–11}. For a number of faults, however, palaeoseismological data have revealed a more complex behaviour, with pronounced temporal variations in earthquake occurrence and fault slip rate over several earthquake cycles^{12,13}. Such highly variable rates of faulting have been incorporated in models that combine earthquake clustering on short timescales with uniform long-term strain accumulation¹⁴, but the specific causes of the slip rate variations remain obscure. Here we use finite-element models to investigate why four parallel normal faults—the Wasatch, West Valley, Oquirrh and Stansbury faults in the eastern Basin and Range Province, Utah—have experienced a simultaneous increase in slip rate during the Late Pleistocene/Holocene³. Faster slip accumulation during the Holocene is best documented for the Wasatch fault, for which geologic^{15,16}, palaeoseismological^{1,2,17} and geodetic¹⁸ data constrain an increase in slip rate by a factor of ~ 3 as compared to the long-term rate determined over timescales of 10^5 to 10^7 yr (ref. 3). Accelerated accumulation of slip on these faults has been interpreted as a clustering of earthquakes during the Holocene^{1–3,19}. Given that the extension rate across the Basin and Range Province has decreased by about 50% since 10 Myr ago^{20,21}, it is not obvious why several parallel faults should experience a synchronous clustering of earthquakes over the past ~ 10 kyr without external forcing, that is, without the addition or the removal of mass from Earth's surface.

The intriguing coincidence between the onset of faster slip accumulation on normal faults in the Wasatch region and the end of the last ice age strongly suggests an external forcing. Invoking postglacial unloading as a cause for increased fault slip rates is, however, difficult to reconcile with the stress state of normal faulting²². Normal faults develop with the maximum principal stress, σ_1 , perpendicular to Earth's surface, and the minimum principal stress, σ_3 , horizontal and perpendicular to the fault's strike. Hence unloading should decrease σ_1 and decelerate fault motion²². Conversely, adding a surface load increases σ_1 and makes the fault more susceptible to failure, which will promote earthquakes and ultimately lead to faster slip.

To evaluate the predicted behaviour of a normal fault during glacial loading and subsequent unloading, we designed a 1,000-km-long finite-element model²³, which incorporates a discrete normal fault²⁴ in a rheologically layered lithosphere. The model geometry and the properties of the fault, crust and lithospheric mantle are shown in Fig. 1. Extension of the model lithosphere at a constant rate leads to a steady fault slip rate of 0.20 mm yr^{-1} (Fig. 2a). In the first experiment, we performed two model runs, one without loading and one in which a spatially uniform load, equivalent to 500 m of ice, is applied on top of the model. The bottom of the model is constrained to move only in the horizontal direction, that is, no isostatic or flexural compensation of the load is allowed. The temporal changes in the load are described by a loading function consisting of three increments: a linear increase to full magnitude within 20 kyr, a constant value for 100 kyr and removal within 10 kyr (Fig. 2b). After unloading, the model runs for an additional 100 kyr. The first experiment demonstrates that during the application of the load the slip rate increases by a factor of ~ 2.5 to 0.49 mm yr^{-1} after a lag of 4 kyr (Fig. 2b). In the presence of the entire load, the slip rate is 0.24 mm yr^{-1} (Fig. 2b), that is, 20% higher than in the control run. The start of unloading stops slip accumulation for 16 kyr, until 6 kyr after complete removal of the

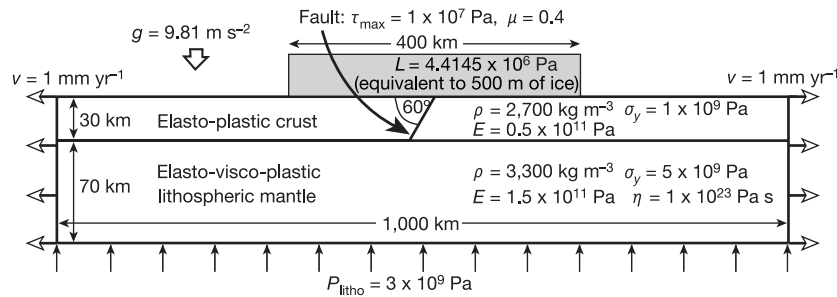


Figure 1 Set-up of the finite element model with a discrete normal fault. Slip on the fault is controlled by a Coulomb failure criterion ($\tau_{\max}$, critical shear stress of fault; μ , coefficient of friction). Other parameters are density (ρ), elastic modulus (E), plastic yield stress (σ_y),

viscosity (η), acceleration due to gravity (g), load (L), velocity (v) and lithostatic pressure (P_{litho}).

load, when the slip rate increases again to 0.19 mm yr^{-1} (Fig. 2b). The modelled changes in slip rate confirm the theoretical prediction of the response of normal faults to the application and removal of a load²². These findings imply that, in absence of isostatic rebound, postglacial unloading cannot explain an increase in the slip rate of a normal fault.

But the first experiment neglects the fact that any fault is embedded in a lithosphere that, when loaded, responds by flexure²⁵. The bending stresses resulting from flexure increase the horizontal stresses in the upper part of the lithosphere, whereas removal of the load causes isostatic rebound, which relaxes the bending stresses and reduces the horizontal stresses^{25–27}. In regions of normal faulting, where σ_1 is vertical and σ_3 is horizontal, lithospheric flexure increases σ_3 and may thus suppress faulting. In contrast, isostatic rebound decreases σ_3 such that faulting may be promoted, despite the removal of mass at the surface. In reality, the temporal changes of the stress state are, however, more complicated as a result of ongoing extension, lateral flow of mantle material, and buoyancy forces generated by loading. For stable continental shields affected by the melting of large ice sheets, the stress changes caused by the interplay between these processes have been evaluated numerically^{26,28}. The resulting variations in the style of faulting have been predicted on a continental scale^{26,28}, but the response of a discrete

fault to lithospheric flexure and rebound has never been quantified.

Our second experiment investigates whether rebound-induced changes in the state of stress may overcome the effect of pure unloading and promote accelerated normal fault slip. To implement lithospheric rebound, we add an elastic foundation and linear dashpot elements to the bottom of the lithosphere, which together mimic the asthenosphere (density $\rho = 3,200 \text{ kg m}^{-3}$, viscosity $\eta = 5 \times 10^{19} \text{ Pa s}$). The bottom of the model is now free to move in both the vertical and horizontal directions. All other parameters remain the same as in the first experiment. The second experiment reveals that, in contrast to the first, slip accumulation ceases $\sim 7.5 \text{ kyr}$ after the onset of loading (Fig. 2c, d). It is not until 12 kyr after the load has reached its full magnitude that the slip rate increases to 0.2 mm yr^{-1} . During unloading the fault experiences a marked slip rate increase to 0.68 mm yr^{-1} , and this high rate prevails for nearly 10 kyr before it drops to a rate that is still slightly higher than that in the control run (Fig. 2d). Hence lithospheric rebound is capable of increasing the slip rate of a normal fault, thereby outweighing the decelerating effect of pure unloading.

We now test whether the Late Pleistocene/Holocene increase in slip rates of the Wasatch, West Valley, Oquirrh and Stansbury normal faults may be due to isostatic rebound caused by the disappearance of Pleistocene Lake Bonneville, of which today's

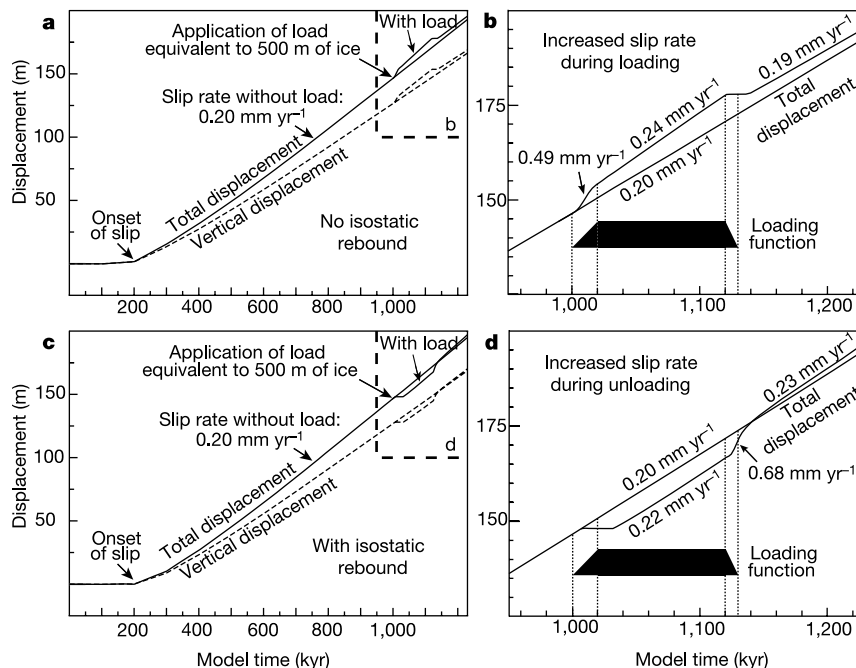


Figure 2 Response of the model normal fault to loading and unloading. **a, b**, Without isostatic rebound; **c, d**, including rebound. Panels **a** and **c** show total and vertical displacement as a function of time; **b** and **d** are enlarged parts of the upper right corner of

a and **c**, respectively. **a, b**, During loading the slip rate is increased, whereas slip accumulation stops during unloading. **c, d**, Slip accumulation stops during loading, but is considerably enhanced during unloading. See text for further explanation.

Great Salt Lake is a tiny remnant. Lake Bonneville covered an area of $\sim 50,000 \text{ km}^2$ east of the Wasatch range and attained a maximum water depth of 350 m at 18–17 kyr ago, before the lake level dropped by $\sim 275 \text{ m}$ within $\sim 4 \text{ kyr}$ (ref. 6). (Radiocarbon ages⁶ have been converted to calendar years²⁹.) Palaeo-shorelines on former islands in Lake Bonneville are located several tens of metres above palaeo-shorelines along its periphery, which demonstrates that isostatic rebound has occurred after regression of the lake^{4–6}. Note that the four faults are located in the centre of the region affected by isostatic rebound⁴, and strike subparallel to the eastern margin of Lake Bonneville. In addition to the lake, valley glaciers in the ranges adjacent to the faults contributed to loading of the Wasatch area^{7,8}. The size and distribution of moraines indicate that these glaciers were probably up to a few hundred metres thick and extended locally across the faults but not into the adjacent basin occupied by Lake Bonneville^{7,8}.

To investigate the effect of Lake Bonneville on the faults in the Wasatch region, we modified the second model by reducing the lithospheric thickness to 60 km (ref. 30) and reducing the viscosity of the asthenosphere to $2 \times 10^{19} \text{ Pa s}$. The load is now distributed asymmetrically across the fault to take into account the fact that loading of the hanging wall of the Wasatch fault resulted from the 200-km-wide Lake Bonneville, whereas footwall loading is caused by glaciers in the Wasatch Range^{7,8} and the Uinta Mountains⁸ farther east. The thickness of these glaciers is not well known. In the model, we use a 100-km-wide and 150-m-thick layer of ice as load on the footwall. The loading function is adjusted to the lake level curve of Lake Bonneville^{6,29}: the lake starts to grow 34 kyr ago and reaches a maximum depth of 350 m at 18–17 kyr ago, before the lake level declines linearly to zero between 17 and 13 kyr ago (Fig. 3). This loading function is also applied to the footwall, as ¹⁴C dating indicates that the growth and decay of glaciers in the Wasatch Mountains⁷ occurred more or less in phase with the expansion and shrinkage of Lake Bonneville. This third model predicts that $\sim 10 \text{ kyr}$ after the onset of loading, that is, at 24 kyr ago, fault slip ceases for 11.5 kyr (Fig. 3). Then 0.5 kyr after complete disappearance of the lake, that is, at 12.5 kyr ago, slip accumulation starts again at a high rate of 0.73 mm yr^{-1} , which gradually decreases to 0.23 mm yr^{-1} (Fig. 3). These model results compare favourably with the slip history of the faults in the Wasatch region, which experienced a period of reduced seismicity during the lifetime of Lake Bonneville³. The marked slip rate increase of the model at 12.5 kyr ago is equivalent to the onset of earthquake clustering in the Late

Pleistocene/Early Holocene, which is best constrained for the Wasatch fault^{1–3}. Whereas the average vertical displacement rate of the Wasatch fault during the Late Pleistocene was $0.2\text{--}0.7 \text{ mm yr}^{-1}$, the Holocene rate is $\sim 1.7 \text{ mm yr}^{-1}$ (refs 1–3). A similar contrast in the integrated slip rates is also observed in our model. The mean vertical slip rate of the model fault over a period of 50 kyr before the onset of slip acceleration is 0.15 mm yr^{-1} , whereas the average rate since 12.5 kyr ago is 0.35 mm yr^{-1} . In conclusion, our third model suggests that lithospheric rebound caused by removal of Lake Bonneville has triggered the marked increase in the slip rates of the faults in the Wasatch region, and is thus responsible for the previously enigmatic clustering of earthquakes during the Holocene.

We have shown that loads applied on Earth's surface exert a fundamental control on the rate of strain accumulation on individual normal faults. Abrupt changes in rates of faulting may occur with a time lag of several thousand years after a load has been applied or removed. Subsequently, the accelerated or decelerated motion of the fault may be sustained over periods of several thousand years. If the external forcing is due to glaciers or lakes, the fault slip history is ultimately related to past climate changes. Reconstructing the Late Quaternary loading histories of active faults by exploring continental climate archives will be essential in order to better understand past and future strain release patterns of seismogenic faults and to evaluate their seismic hazard. □

Received 15 October 2004; accepted 10 March 2005; doi:10.1038/nature03562.

1. Machette, M., Personius, S. F. & Nelson, A. R. The Wasatch fault zone, Utah—segmentation and history of Holocene earthquakes. *J. Struct. Geol.* **13**, 137–149 (1991).
2. McCalpin, J. P. & Nishenko, S. P. Holocene paleoseismicity, temporal clustering, and probabilities of future large ($M > 7$) earthquakes on the Wasatch fault zone, Utah. *J. Geophys. Res.* **101**, 6233–6253 (1996).
3. Friedrich, A., Wernicke, B. P., Niemi, N. A., Bennett, R. A. & Davis, J. L. Comparison of geodetic and geologic data from the Wasatch region, Utah, and implications for the spectral character of Earth deformation at periods of 10 to 10 million years. *J. Geophys. Res.* **108**, 2199 doi:10.1029/2001JB000682 (2003).
4. Bills, B. G., Currey, D. R. & Marshall, G. A. Viscosity estimates for the crust and upper mantle from patterns of lacustrine shoreline deformation in the Eastern Great Basin. *J. Geophys. Res.* **99** (B11), 22059–22086 (1994).
5. Bills, B. G. & May, G. M. Constraints on lithospheric thickness and upper mantle viscosity from isostatic warping of Bonneville, Provo, and Gilbert stage shorelines. *J. Geophys. Res.* **92**, 11493–11508 (1987).
6. Oviatt, C. G., Currey, D. R. & Sack, D. Radiocarbon chronology of Lake Bonneville, Eastern Great Basin, USA. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **99**, 225–241 (1992).
7. Madsen, D. B. & Currey, D. R. Late Quaternary glacial and vegetation changes, Little Cottonwood Canyon, Wasatch Mountains, Utah. *Quat. Res.* **12**, 254–270 (1979).
8. Osborn, G. & Bevis, K. Glaciation in the Great Basin of the Western United States. *Quat. Sci. Rev.* **20**, 1377–1410 (2001).
9. Reid, H. F. *Report of the (California) State Earthquake Investigation Commission* 1–192 (Publ. 87, Vol. 2, Calif. State Earthquake Invest. Comm., Sacramento, 1910).
10. Shimazaki, K. & Nakata, T. Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett.* **7**, 279–282 (1980).
11. Thatcher, W. The earthquake deformation cycle, recurrence, and the time-predictable model. *J. Geophys. Res.* **89**, 5674–5680 (1984).
12. Grant, L. & Sieh, K. Paleoseismic evidence of clustered earthquakes on the San Andreas fault in the Carrizo Plain, California. *J. Geophys. Res.* **99**, 6819–6841 (1994).
13. Marco, S., Stein, M., Agnon, A. & Ron, H. Long-term earthquake clustering: A 5000-year paleoseismic record in the Dead Sea Graben. *J. Geophys. Res.* **101**, 6179–6191 (1996).
14. Wallace, R. E. Grouping and migration of surface faulting and variation in slip rates on faults in the Great Basin province. *Bull. Seismol. Soc. Am.* **77**, 868–877 (1987).
15. Parry, W. T. & Bruhn, R. L. Pore fluid and seismogenic characteristics of the fault rock at depth on the Wasatch Fault, Utah. *J. Geophys. Res.* **91**, 730–744 (1986).
16. Ehlers, T. A., Willett, S. D., Armstrong, P. A. & Chapman, D. S. Exhumation of the central Wasatch Mountains, Utah: 2. Thermokinematic model of exhumation, erosion, and thermochronometer interpretation. *J. Geophys. Res.* **108**, 2173 doi:10.1029/2001JB001723 (2003).
17. Schwartz, D. P. & Coppersmith, K. J. Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones. *J. Geophys. Res.* **89**, 5681–5698 (1984).
18. Martinez, L. J., Meertens, C. M. & Smith, R. B. Rapid deformation rates along the Wasatch fault zone, Utah, from first GPS measurements with implications for earthquake hazard. *Geophys. Res. Lett.* **25**, 567–570 (1998).
19. Malservisi, R., Dixon, T. H., La Femina, P. C. & Furlong, K. P. Holocene slip rates of the Wasatch fault zone, Utah, from geodetic data: Earthquake cycle effects. *Geophys. Res. Lett.* **30**, 1673 doi:10.1029/2003GL017408 (2003).
20. Wernicke, B., Axen, G. J. & Snow, J. K. Basin and Range extensional tectonics at the latitude of Las Vegas, Nevada. *Geol. Soc. Am. Bull.* **100**, 1738–1757 (1988).
21. Wernicke, B. & Snow, J. K. Cenozoic tectonism in the central Basin and Range: Motion of the Sierran-Great Valley Block. *Int. Geol. Rev.* **40**, 403–410 (1998).
22. Johnston, A. C. in *Earthquakes at North-Atlantic Passive Margins: Neotectonics and Postglacial Rebound* (eds Gregersen, S. & Basham, P. W.) 581–599 (Kluwer, Dordrecht, 1989).

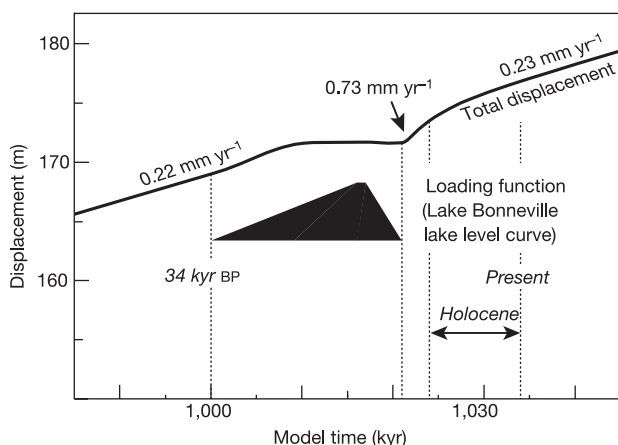


Figure 3 Results of a model simulating normal faulting in the Wasatch region. The loading function corresponds to the calibrated²⁹ lake level curve of Lake Bonneville⁶. Note that the horizontal scale differs from Fig. 2. Loading of the fault causes fault motion to stop at $\sim 24 \text{ kyr}$ ago. At $\sim 12.5 \text{ kyr}$ ago, lithospheric rebound following unloading causes a sudden increase of the slip rate to 0.73 mm yr^{-1} . See text for further explanation. BP, before present.

23. ABAQUS/Standard, User's Manual, version 6.3 (Hibbitt, Karlson and Sorenson Inc., Pawtucket, Rhode Island, 2002).
24. Ellis, S. & Stöckhert, B. Elevated stresses and creep rates beneath the brittle-ductile transition caused by seismic faulting in the upper crust. *J. Geophys. Res.* **109**, B05407, doi:10.1029/2003JB002744 (2004).
25. Watts, A. B. *Isostasy and Flexure of the Lithosphere* (Cambridge Univ. Press, Cambridge, UK, 2001).
26. Wu, P. & Hasegawa, H. S. Induced stresses and fault potential in eastern Canada due to a disc load: a preliminary analysis. *Geophys. J. Int.* **125**, 415–430 (1996).
27. Stewart, S., Sauber, J. & Rose, J. Glacio-seismotectonics: ice sheets, crustal deformation and seismicity. *Quat. Sci. Rev.* **19**, 1367–1389 (2000).
28. Wu, P., Johnston, P. & Lambeck, K. Postglacial rebound and fault instability in Fennoscandia. *Geophys. J. Int.* **139**, 657–670 (1999).
29. Hughen, K. *et al.* ^{14}C activity and global carbon cycle changes over the past 50,000 years. *Science* **303**, 202–207 (2004).
30. Smith, R. B. *et al.* Geophysical and tectonic framework of the eastern Basin and Range–Colorado Plateau–Rocky Mountains transition. *Geol. Soc. Am. Mem.* **172**, 205–233 (1989).

Acknowledgements We thank A. Friedrich and A. Densmore for discussions. R.H. was supported by a Heisenberg fellowship from the German Research Foundation (DFG).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.H. (rahetzelt@uni-muenster.de).

A primitive therizinosauroid dinosaur from the Early Cretaceous of Utah

James I. Kirkland¹, Lindsay E. Zanno², Scott D. Sampson²,
James M. Clark³ & Donald D. DeBlieux¹

¹Utah Geological Survey, PO Box 146100, Salt Lake City, Utah 84114-6100, USA

²Utah Museum of Natural History and Department of Geology and Geophysics, University of Utah, 1390 E. Presidents Circle, Salt Lake City, Utah 84112-0050, USA

³George Washington University, Washington DC 20052, USA

Therizinosauroids are an enigmatic group of dinosaurs known mostly from the Cretaceous period of Asia, whose derived members are characterized by elongate necks, laterally expanded pelves, small, leaf-shaped teeth, edentulous rostra and mandibular symphyses that probably bore keratinized beaks^{1,2}. Although more than a dozen therizinosauroid taxa are known, their relationships within Dinosauria have remained controversial because of fragmentary remains and an unusual suite of characters. The recently discovered 'feathered' therizinosauroid *Beipiaosaurus* from the Early Cretaceous of China helped to clarify the theropod affinities of the group³. However, *Beipiaosaurus* is also poorly represented. Here we describe a new, primitive therizinosauroid from an extensive paucispecific bonebed at the base of the Cedar Mountain Formation (Early Cretaceous) of east-central Utah^{4,5}. This new taxon represents the most complete and most basal therizinosauroid yet discovered. Phylogenetic analysis of coelurosaurian theropods incorporating this taxon places it at the base of the clade Therizinosauroiden, indicating that this species documents the earliest known stage in the poorly understood transition from carnivory to herbivory within Therizinosauroida. The taxon provides the first documentation, to our knowledge, of therizinosauroids in North America during the Early Cretaceous.

Theropoda Marsh 1881

Coelurosauria von Huene 1914

Therizinosauroida Maleev 1954

Falcarius utahensis gen. et sp. nov.

Etymology. From *Falcarius* (Latin, a sickle-maker) and *utahensis* (refers to Utah as its place of origin).

Holotype. Utah Museum of Natural History; UMNH VP 15000 partial braincase (Fig. 1b, c).

Referred specimens (paratypes). Utah Museum of Natural History; UMNH VP 12279–12443, 14524–14999, 15001–15149.

Horizon and locality. The Crystal Geyser Quarry, Grand County, Utah, lies at the base of the Cedar Mountain Formation, directly but disconformably overlying the Upper Jurassic Morrison Formation (UMNH VP locality no. 157). The basal Yellow Cat Member is estimated to be Barremian in age on the basis of the preserved dinosaur fauna (the polacanthine ankylosaur *Gastonia*, the ornithomimid *Iguanodon*, three species of sauropods, and large and small theropods such as *Utahraptor* and *Nedcolbertia*^{4,5}), charophytes⁴ and palynomorphs (D. Eberth and B. Britt, personal communication 2004).

Diagnosis. *Falcarius* is a maniraptoran theropod diagnosed by the following suite of unique characters from holotype braincase: expansive, open pneumatic cavities in basioccipital and basisphenoid with the basisphenoidal recess directed ventrally; and extensive subcondylar pneumatic recesses lateral to the occipital condyle. From paratype elements: presence of one enlarged tooth at anterior-most end of dentary; laterally deflected and biconcave apex of the deltopectoral crest; posterior tuberosity on the distal end of the humerus opposite the radial condyle; flexor tubercle of phalanx I–II with well-defined collateral pits on the distal aspect; and the unique combination of the following features: dentary lacking lateral shelf found in other therizinosauroids; mid-caudal vertebrae elongate with zygapophyses less than one-third the length of centrum as opposed to shortened vertebrae in other therizinosauroids and oviraptorosaurs.

Description. *Falcarius* is a gracile, small- to medium-sized theropod, approximately 1 m in height at hips and 4 m in length. The specimens referred to this taxon account for approximately 90% of the skeleton (Figs 1 and 2).

Skull material is underrepresented and currently includes a left maxilla, right postorbital, paired frontals, two braincases, isolated juvenile basiocciput, left and right quadrates, one left and two right dentaries, and a right splenial. As in *Erlikosaurus*⁶, the number of maxillary neurovascular foramina is reduced to approximately one slit-like foramen for every three alveoli. The quadrates are anteriorly arched with large pneumatic fossae. The frontals are subrectangular, as in dromaeosaurids and oviraptorosaurs, with similarly inflated cerebral fossa (Fig. 1). The fossa for the olfactory bulb is approximately two-thirds as wide as the cerebral fossa. The paroccipital processes are more slender and elongate than in derived therizinosauroids and are pneumatized via caudal tympanic recesses (Fig. 1). The basioccipital is pocketed ventrolaterally by extensive subcondylar recesses that encompass two openings for cranial nerves X–XI and a single opening for XII. *Falcarius* lacks a dorsoventral expansion of the paroccipital process and a ventral pneumatic inflation of the basisphenoid, the derived condition found in both *Erlikosaurus*⁶ and *Nothronychus*⁷.

The dentary deepens posteriorly, but lacks the lateral shelf and down-turned symphyseal region possessed by all other known therizinosauroids^{2,3,6} (Fig. 2). The dentary is medially curved anteriorly, indicating the presence of a U-shaped snout as in other therizinosauroids^{2,6} and oviraptorosaurs⁸.

The dentary teeth share several features with the teeth of other therizinosauroids (Fig. 2). Similarities include posteriorly small, lanceolate and basally constricted crowns that become taller anteriorly, as well as the presence of inflated, circular roots. Also, in contrast to other therizinosauroids, conical teeth are present in the anteriormost portion of the dentary in *Falcarius*. The first alveolus of the dentary is hypertrophied and apparently housed a tooth with twice the cross-sectional area of more posterior tooth positions. This condition is comparable to the enlarged premaxillary teeth of the primitive oviraptorosaur *Incisivosaurus* (= *Protoarchaeopteryx*⁹), although its anterior dentary is

edentulous¹⁰. The preserved teeth of the fragmentary maxilla have lower tooth crowns that decrease in size posteriorly (Fig. 2). The denticles on all teeth are small (7–10 mm⁻¹), unlike other therizinosauroids (Fig. 2).

The vertebral column is well represented with the exception of the axis. Anterior cervical centra are amphicoelous, elongate (length four times height) and highly pneumatized, with multiple, small pneumatic foramina. As in derived therizinosauroids^{2,11–13} and oviraptorosaurs¹⁴, cervical neural arches are low and elongate and zygapophyses become more widely spaced posteriorly (giving the neural arch an X-shape in dorsal view). Dorsal neural arches have large transverse processes supported by multiple laminae, short neural spines that thicken dorsally, and thick hyposphene–hypantrum articulations. There are five sacral vertebrae as in *Alxasaurus*¹², in contrast to six preserved in *Neimongosaurus*¹⁵ and *Segnosaurus*¹⁶, with the intermediate three possessing pneumatic fossae. Medial to distal caudals are more than four times as long as high. All caudal vertebrae seem to be apneumatic with loss of caudal ribs and neural spines distal to caudals 11–13 (Fig. 1), in contrast to

the condition in derived oviraptorosaurs¹³ and derived therizinosauroids^{2,13}.

The coracoids are strongly recurved with a pronounced ventral tubercle. As in more derived therizinosauroids, the humeri have an angular internal tuberosity, cranially positioned distal condyles and a hypertrophied entepicondyle^{1,2}. The humeral shaft lacks both the posterior tuberosity of more derived therizinosauroids^{1,2,16–19} and the anterior tuberosity proximal to the entepicondyle of *Neimongosaurus*¹⁵ and *Erliansaurus*²⁰. The ulna is bowed with a robust olecranon process. The radius is straight and lacks the biceps tubercle seen on *Neimongosaurus*¹⁵. The manus is relatively gracile and elongate, as in *Beipiaosaurus*³ and many oviraptorosaurs¹⁴ (Fig. 1). Distal carpals are preserved in two morphs: one fused into a semilunate that caps metacarpals I and II, and the other as unfused pairs of distal carpals as in other therizinosauroids². Metacarpal I exhibits a rectangular buttress on its proximoventral surface, similar to that of *Alxasaurus*¹².

The pelvis is primitive in that the ilium is relatively elongate anteroposteriorly, with a parasagittal dorsal margin lateral to the

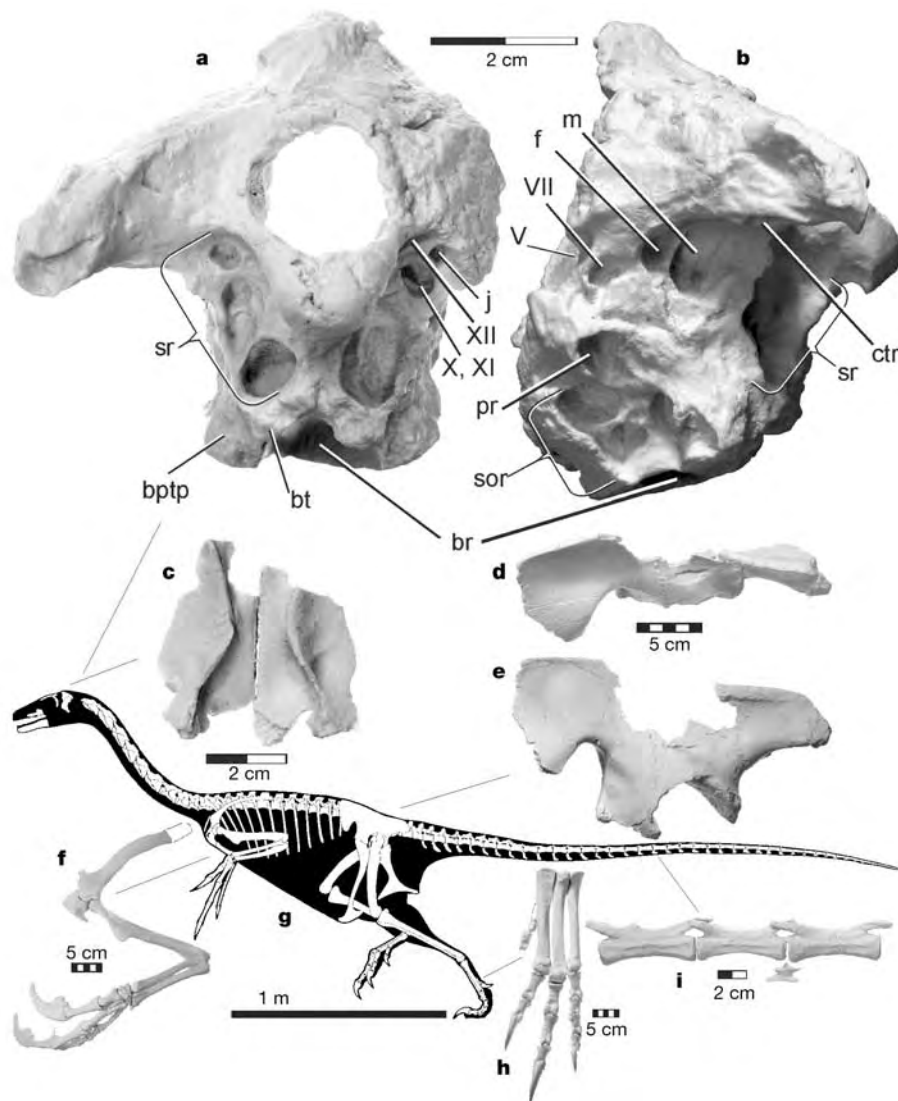


Figure 1 Skeletal elements of *Falcarius utahensis*. **a, b**, Holotype braincase UMNH-VP 15000 in posterior (**a**) and lateral (**b**) views; **c**, paired frontals UMNH VP 14524, 14525 in ventral view; **d, e**, left ilium UMNH VP12368 in dorsal (**d**) and lateral (**e**) views; **f**, reconstructed right pectoral girdle UMNH VP 12279, 12281 and forelimb UMNH VP 12284, 12287, 12289, 12291, 12294–12296, 12298, 12300, 12302, 12314, 12304, 12306, 12308, 12310, 12312, 12316, 12320 in lateral view; **g**, *Falcarius* skeletal

reconstruction; **h**, reconstructed left foot UMNH-VP 12330–12342, 12352–12356 in anterioposterior view; **i**, medial caudal vertebrae UMNH VP 12405, 12407, 12409 with chevron UMNH VP12391 in lateral view. bptp, basiptyergoid process; br, basisphenoidal recess; bt, basitubera; ctr, caudal tympanic recess; f, fenestra ovalis and fenestra pseudorotunda; j, jugular foramina; m, metotic fissure; pr, prootic recess; sor, subotic recess; sr, subcondylar recess; roman numerals represent cranial nerves.

sacral transverse processes (Fig. 1), the pubis has not rotated to an opisthopubic condition, and the obturator process does not contact the pubis. The pubic peduncle is also primitive, being wide and dorsoventrally shallow. Derived therizinosauroid characters of the pelvis include a deep, strongly hooked and laterally flaring preacetabular process, and a dorsoventrally shortened postacetabular process of the ilium². The acetabulum of *Falcarius* is anteriomedially directed, a condition also present in *Neimongosaurus*¹⁵.

As with the forelimb, the hindlimb of *Falcarius* is relatively elongate and slender, with an anteriorly bowed femur estimated at approximately 85% tibial length. The lesser trochanter is alar, separated from the greater trochanter by a deep, narrow cleft. The tibia is gracile with a pronounced fibular crest. As in other therizinosauroids^{2,3,12,13}, the proximal fibula lacks the medial fossa found on most maniraptorans. The distal tibia is surrounded by the astragalus, unlike the reduced condition in other therizinosauroids². The ascending process of the astragalus is tall and nearly symmetrical, unlike the shortened, asymmetrical morphology of *Therizinosaur*²¹. The functionally tridactyl pes is similar to that of *Beipiaosaurus*³ and oviraptorosaurs¹⁴ and unlike the functionally tetradactyl pes of all other known therizinosauroids^{1,2} (Fig. 1).

A phylogenetic analysis (Fig. 3) provides strong support for the hypothesis that *Falcarius* is the basalmost therizinosauroid known. *Falcarius* possesses several maniraptoran synapomorphies not previously documented in basal therizinosauroids, including hypapophyses on the presacral vertebral column, bowed ulna, presence of a semilunate comprised of fused distal carpals one and two, ischium less than two-thirds pubis length, and distally positioned obturator process of the ischium^{22–25}.

Therizinosauroids are here proposed as shifting their dietary habit from predation to herbivory on the basis of the development of a number of features that seem convergent with clades of other herbivorous dinosaurs. The most significant of these features include small, leaf-shaped teeth, an edentulous beak, posterior displacement of the pubis and lateral expansion of the pelvis associated with greatly increased intestinal volume^{1,26}, and shortening of the tibia relative to the femur and an increased number of weight-supporting pedal digits—the latter two being specific reversals of the cursorial

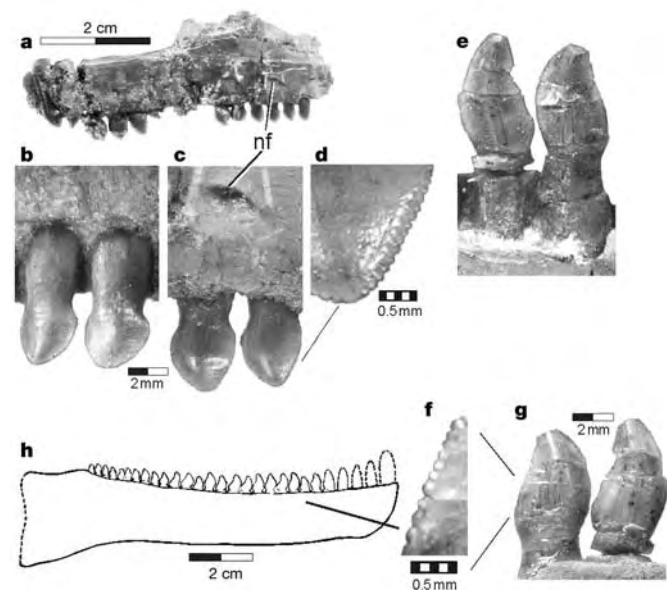


Figure 2 Jaws and teeth of *Falcarius utahensis*. **a**, Partial left maxilla UMNH VP 14526 in lateral view; **b**, **c**, close-up of maxillary teeth in lingual (**b**) and labial (**c**) views; **d**, detail of maxillary tooth denticles in labial view; **e**, dentary teeth UMNH VP 14527 in lingual view; **f**, detail of dentary tooth denticles in labial view; **g**, same dentary teeth in labial view; **h**, composite reconstruction of the dentary of *Falcarius* based on UMNH VP 14527, UMNH VP 14528 and UMNH VP 14529. nf, nutritive foramina.

condition. *Falcarius* demonstrates the mosaic nature of this evolutionary transition, indicating that the dentition and pelvis were among the first hard-tissue structures to undergo modification. These changes probably coincided with modifications in food acquisition and digestion during the early stages of therizinosauroid

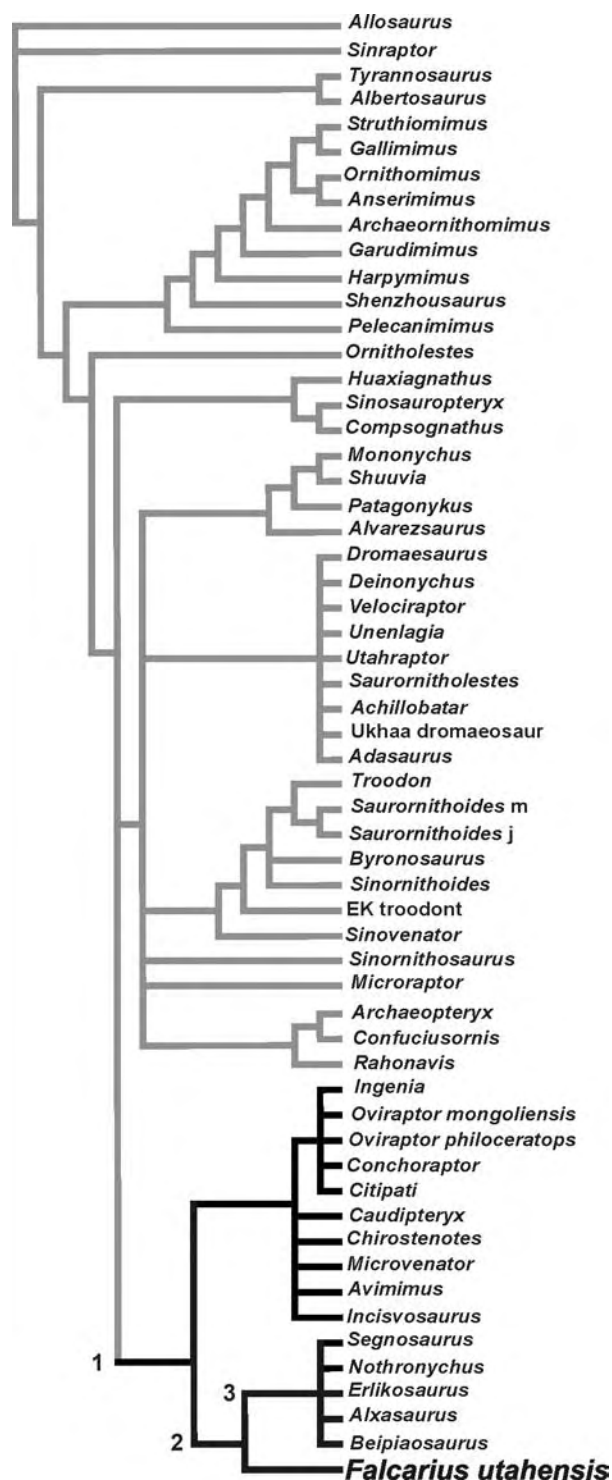


Figure 3 Phylogenetic relationships of *Falcarius* among the Coelurosauria. Strict consensus of 5,000 trees with the following statistics: treelength 679; Consistency Index, 0.41; Retention Index, 0.74; Rescaled Consistency Index, 0.31. Therizinosauroidae + Oviraptorosauria (node 1), Therizinosauroidae (node 2): Therizinosauroidae more derived than *Falcarius* (node 3). Character states defining these nodes are discussed in Methods. Advent of dietary shift from carnivory to herbivory is postulated to occur at base clade 1 (shaded black).

evolution. Moreover, similarities between the dentition of the basal therizinosaur *Falcarius* and the basal oviraptorosaur *Incisivosaurus*, in combination with their proposed sister relationship (Fig. 3), raises the possibility that the common ancestor of these clades had already undertaken the initial steps in this transition.

The Early Jurassic Chinese jaw from the Lufeng Series in southern China described tentatively as the basal therizinosaurid *Eshanosaurus*²⁷ is more derived in having a lateral shelf and down-turned symphysis, which are absent in *Falcarius*. Its therizinosaurid identification have been considered problematic by some authors^{13,25}. *Falcarius* casts further doubt on the affinities of *Eshanosaurus* by increasing its stratigraphic and phylogenetic inconsistency. Given the discovery of North American members of the therizinosaurid clade, together with the poor record of Middle Cretaceous dinosaurs, it seems that the generally accepted hypothesis of an Asian origin and radiation for Therizinosauridae^{1,2} requires additional testing. □

Methods

Collections from the Crystal Geyser Quarry made so far preserve a minimum of ten individuals based on prepared femora, but the quarry size indicates that perhaps hundreds of disarticulated individuals remain interred, representing multiple growth stages as well as robust and gracile morphotypes. The locality spans about 8,000 m² and the bone-bearing stratum is, on average, 1 m thick, with bone densities in some areas exceeding 100 elements per cubic metre. There is no evidence of another small theropod taxon, so all of the therizinosaurid materials are here referred to *Falcarius*. In addition to the therizinosaurid elements, the quarry contains rare remains of an unidentified ankylosaur. About 99% of all identified bones from the type locality (about 2,000 identifiable bones) represent *Falcarius*.

Several features of these fossils, including neurocentral fusion and fusion of cervical ribs to the vertebrae, indicate that the largest elements represent adult or near-adult individuals. Dorsal ribs and gastralia are the most poorly represented with only two dorsal ribs and one gastralia identified, although many bone fragments in the quarry are thought to be unidentifiable pieces of rib. There are multiple examples of nearly all bones, although much excavation and preparation remains before a detailed taphonomic analysis can be attempted. Overlapping examples of the largest elements are close to the same size, indicating that mature animals had restricted growth. Additionally, associated elements (particularly appendicular elements) indicate the relative proportions of skeletal elements in some individuals. The skeletal reconstruction of *Falcarius* in Fig. 1 was based on the largest preserved elements in the quarry, with vertebral numbers estimated on changing proportions of the vertebrae and comparisons with other maniraptoran theropods.

Phylogenetic relationships were analysed with the use of data from published sources supplemented with novel information (see Supplementary Information). Parsimony analysis of 57 taxa and 231 characters was performed with PAUP 4.0b10 (ref. 28), with all characters weighted equally and a single character ordered, using *Allosaurus* and *Sinraptor* as outgroups (Fig. 3).

Five unambiguous synapomorphies support Therizinosauridae: teeth serrated (reversal); ventral surface of dentary descends strongly posteriorly; ventral edge of anterior ala of ilium hooked anteriorly; distal humerus with large medial condyle, centred on distal end; and preacetabular portion of ilium laterally flaring. The basal position of *Falcarius* is supported by the absence of the following characters: labial face of dentary with lateral ridge and inset tooth row; interdental plates on dentary; obturator process of ischium does not contact pubis; metatarsal I without proximal articulating surface. Therizinosauridae + Oviraptorosauria characterized by basiterygoid processes abbreviated or absent; suborbital fenestra reduced in size or absent; basiterygoid processes hollow; symphyseal region of dentary strongly recurved; maxillary and dentary teeth lanceolate and subsymmetrical.

Received 10 December 2004; accepted 17 February 2005; doi:10.1038/nature03468.

- Maryanska, T. in *The Complete Dinosaur* (eds Farlow, J. O. & Brett-Surman, M. K.) 234–241 (Indiana Univ. Press, Bloomington, 1997).
- Clark, J. M., Maryanska, T. & Barsbold, R. in *The Dinosauria II* (eds Weishampel, D. B., Dodson, P. & Osmolska, H.) 151–164 (Univ. California Press, Berkeley, 2004).
- Xu, X., Tang, Z.-L. & Wang, X.-L. A therizinosaurid dinosaur with integumentary structures from China. *Nature* **399**, 350–354 (1999).
- Kirkland, J. I., et al. in *Vertebrate Paleontology in Utah* (ed. Gillette, D.) 201–217 (Utah Geol. Surv. Misc. Pub. 99-1, Salt Lake City, 1999).
- Kirkland, J. I. Utah's newly recognized dinosaur record from the Early Cretaceous Cedar Mountain Formation. *Utah Geol. Surv. Notes* **37**, 1–5 (2005).
- Clark, J. M., Perle, A. & Norell, M. A. The skull of *Erlicosaurus andrewsi*, a Late Cretaceous 'segnosaur' (Theropoda: Therizinosauridae) from Mongolia. *Am. Mus. Novit.* **3115**, 1–39 (1994).
- Kirkland, J. I. & Wolfe, D. G. A therizinosaurid (Dinosauria: Theropoda) braincase from the middle Turonian (Cretaceous) of North America. *J. Vertebr. Paleont.* **21**, 68A (2001).
- Sues, H.-D. On *Chirostenotes*, a Late Cretaceous oviraptorosaur (Dinosauria: Theropoda) from western North America. *J. Vertebr. Paleont.* **17**, 698–716 (1997).
- Senter, P., Barsbold, R., Britt, B. B. & Burnham, D. A. Systematics and evolution of Dromaeosauridae (Dinosauria, Theropoda). *Bull. Gunma Mus. Nat. Hist.* **8**, 1–20 (2004).
- Xu, X., Cheng, Y.-N., Wang, X.-L. & Chang, C.-H. An unusual oviraptorosaurian dinosaur from China. *Nature* **419**, 291–293 (2002).

- Dong, Z. & Yu, H. in *Sino-Japanese Silk Road Dinosaur Expedition* (ed. Dong, Z.) 90–95 (China Ocean Press, Beijing, 1997).
- Russell, D. A. & Dong, Z.-M. The affinities of a new theropod from the Alxa Desert, Inner Mongolia, People's Republic of China. *Can. J. Earth Sci.* **30**, 2107–2127 (1993).
- Kirkland, J. I. & Wolfe, D. G. First definitive therizinosaurid (Dinosauria; Theropoda) from North America. *J. Vertebr. Paleont.* **21**, 410–414 (2001).
- Osmolska, H., Currie, P. J. & Barsbold, R. in *The Dinosauria II* (eds Weishampel, D. B., Dodson, P. & Osmolska, H.) 165–183 (Univ. California Press, Berkeley, 2004).
- Zhang, X.-H. et al. A long-necked therizinosaurid dinosaur from the Upper Cretaceous Iren Dabasu Formation of Nei Mongol, People's Republic of China. *Vertebr. Palasiatica* **10**, 282–290 (2001).
- Perle, A. Segnosauridae—a new family of theropods from the Late Cretaceous of Mongolia. *Trans. Joint Soviet-Mongolian Paleont. Exped.* **8**, 45–55 (1979).
- Barsbold, R. New information on Therizinosauridae (Theropoda). *Trans. Joint Soviet-Mongolian Paleont. Exped.* **3**, 76–92 (1976).
- Perle, A. A new segnosaurid from the Upper Cretaceous of Mongolia. *Trans. Joint Soviet-Mongolian Paleont. Exped.* **15**, 45–55 (1981).
- Mader, B. J. & Bradley, R. L. A redescription and revised diagnosis of the syntypes of *Alectrosaurus olseni*. *J. Vertebr. Paleont.* **9**, 41–55 (1989).
- Xu, X. et al. A new therizinosaurid (Dinosauria, Theropoda) from the Upper Cretaceous Iren Dabasu Formation of Nei Mongol. *Vertebr. Palasiatica* **40**, 228–240 (2002).
- Perle, A. On a new finding of the hindlimb of *Therizinosaur* sp. from the Late Cretaceous of Mongolia. *Probl. Geol. Mongol.* **5**, 94–98 (1982).
- Gauthier, J. A. Saurischian monophyly and the origin of birds. *Mem. Calif. Acad. Sci.* **8**, 1–55 (1986).
- Holtz, T. R. A new phylogeny of the carnivorous dinosaurs. *Gaia* **15**, 5–61 (2000).
- Norell, M. A., Clark, J. M. & Makovicky, P. J. in *New Perspectives on the Origin and Early Evolution of Birds* (eds Gauthier, J. & Gall, L. F.) 49–67 (Peabody Museum of Natural History, New Haven, 2001).
- Rauhut, O. W. M. The interrelationships and evolution of basal theropod dinosaurs. *Spec. Pap. Paleont.* **69**, 1–213 (2003).
- Paul, G. S. The segnosaurian dinosaurs: relics of the prosauropod–ornithischian transition. *J. Vertebr. Paleont.* **4**, 507–515 (1984).
- Xu, X., Zhao, X. & Clark, J. M. A new therizinosaur from the Lower Jurassic Lufeng Formation of Yunnan, China. *J. Vertebr. Paleont.* **21**, 477–483 (2001).
- Swofford, D. L. *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4.0b10 (Sinauer Associates, Sunderland, Massachusetts, 2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank X. Xu, Z. Dong, P. Sereno and M. Norell for access to reference material; G. Paul for executing the skeletal reconstruction of *Falcarius*; The Utah Friends of Paleontology, the Utah Museum of Natural History, and H. and P. Bollen for assistance in the field and the laboratory; M. Hayden, M. Christopher, M. Suarez and C. Suarez for overseeing the volunteers in the field; J. R. Scandizzo for introductions to the site discoverer L. Walker of Moab, Utah; R. W. Gaston for providing research casts; D. Smith for comments on theropod braincase anatomy; and J. Harris, M. Hayden, M. Lowe and M. Hyland for reviewing the manuscript. Excavations were conducted under a US Bureau of Land Management permit. Funding was provided by the Utah Geological Survey, Discovery Channel Quest Grants, Paleoscene, to J.I.K. and the NSF to J.M.C.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.I.K. (jameskirkland@utah.gov).

Discovery of the first Asian plethodontid salamander

M. S. Min¹, S. Y. Yang², R. M. Bonett³, D. R. Vieites³, R. A. Brandon⁴ & D. B. Wake³

¹Conservation Genome Resource Bank for Korean Wildlife, and Brain Korea 21, School of Agricultural Biotechnology, Seoul National University, Seoul 151-742, South Korea

²Department of Biology, Inha University, Incheon 402-751, South Korea

³Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, California 94720, USA

⁴Department of Zoology, Southern Illinois University, Carbondale, Illinois 62901, USA

Nearly 70% of the 535 species of salamanders in the world are members of a single family, the Plethodontidae, or lungless salamanders¹. The centre of diversity for this clade is North and Middle America, where the vast majority (99%) of species are found. We report the discovery of the first Asian plethodontid

salamander, from montane woodlands in southwestern Korea. The new species superficially resembles members of North American genera, in particular the morphologically conservative genus *Plethodon*. However, phylogenetic analysis of the nuclear encoded gene *Rag-1* shows the new taxon to be widely divergent from *Plethodon*. The new salamander differs osteologically from putative relatives, especially with respect to the tongue (attached protrusible) and the derived tarsus^{2–6}. We place the species in a new genus on the basis of the morphological and molecular data. The distribution of the new salamander adds to the enigma of Old World plethodontids, which are otherwise restricted to the western Mediterranean region^{7,8}, suggesting a more extensive past distribution of the family.

Amphibia L., 1758
Urodela Duméril, 1806
Plethodontidae Gray, 1850

Karsenia koreana gen. et sp. nov.

Etymology. The genus name honours the discoverer, Stephen J. Karsen. The species name refers to the Korean peninsula. The suggested vernacular name is Korean crevice salamander (Korean: Ikkee dorongyong).

Holotype. EWNHM (Ewha Women's University Natural History Museum) 80314, an adult male from the vicinity of Jangtae-san, Daejeon-si, Chungcheongnam-do, Korea (36° 13' N, 127° 20' W), elevation 210 m, in a rocky, forested habitat, collected 12 April 2003 by Karsen.

Paratypes. MVZ (Museum of Vertebrate Zoology) 246033, 247157, Daedun-san, Keoncheon-ri, Nami-myeon, Kumsan-gun, Chungcheongnam-do, Korea (36° 7.4' N, 127° 19.3' E); MVZ 247156 (Fig. 1), Dae-a-ri, Dongsang-myeon, Wanju-gun, Cheollabuk-do, Korea (35° 59' N, 127° 16.4' E); MCZ (Museum of Comparative Zoology)-A-136691, MVZ 247154–247155, MMS (Laboratory of Mi-Sook Min) 27–29, 32–44 Yeomchi, Yeomchi-ri, Munui-myeon, Cheongwon-gun, Chungcheongbuk-do, Korea (35° 32' N, 127° 32' E); SIUC (Southern Illinois University Collections) H7890–7897, MMS 1–3, 9–20, same locality data as holotype.

Diagnosis of genus and species. A plethodontid salamander on the basis of its lunglessness, absence of a pterygoid bone in adults, large patches of paravomerine teeth, and nasolabial groove. Differs from phenotypically similar *Plethodon* in ankle morphology: distal tarsal 5 (d5) is relatively large and articulates with the centrale, while d4 is relatively small and is separated from the fibulare by d5 (Fig. 2). This arrangement is known only in the plethodontids *Aneides* and

Chiropterotriton^{2,3,6}. *Karsenia* differs from both in having paired premaxillaries rather than a single bone, from *Aneides* in having small hands and feet with short, pointed digits rather than long, terminally expanded digits, and from *Chiropterotriton* in having a tongue that is attached to the front of the mouth and epibranchials that are shorter than the ceratobranchials (rather than being much longer).

This is a relatively small, moderately robust species. Adult males are 38.5–47.7 mm snout-to-vent length (svl) (mean $\pm$ s.d. 41.8 mm $\pm$ 3.6 mm, $n = 8$) and adult females, 38.5–47.7 mm svl (mean $\pm$ s.d. 41.8 $\pm$ 3.7 mm, $n = 12$). A well-defined, relatively broad head (0.13–0.17 $\times$ svl (mean 0.15) in males, 0.13–0.16 $\times$ svl (mean 0.14) in females), is delimited from the body by a marked gular fold. The nasolabial groove is shallow, terminating in a small swelling on the lip. Snouts are more rounded in females than in males, and nostrils are separated by approximately 2.5 mm in both. There are 14–15 costal grooves (mean 14.3 in males, 14.7 in females). Limbs and digits are relatively short, and digits are slightly webbed (Fig. 2). The combined limb length is 0.51–0.55 (mean 0.54) $\times$ axilla–groin length in males, 0.52–0.58 (mean 0.54) in females, and appressed limbs leave three to four intercostal folds uncovered. The tail is round basally but laterally compressed more distally in some individuals, tapering to a sharp tip; the tail is partly regenerated in many specimens, but complete tails are about equal in length to the svl (longest tail 1.1 $\times$ svl). Small eyes (horizontal opening 2.6–2.7 mm) are moderately protuberant and separated by 3.3–4 mm. The rounded tongue, attached anteriorly by genio-

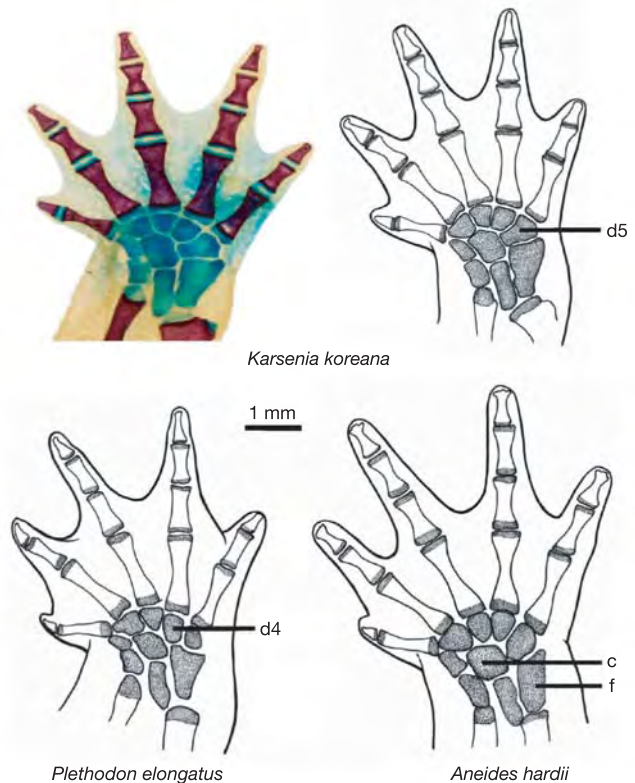


Figure 2 Right foot of three species of plethodontid salamanders. Top, paratype of *Karsenia koreana* (SIUC H7894), stained (alcian blue for cartilage; alizarin red for bone) and cleared in trypsin and glycerine (left) and drawn (right). Bottom, *Plethodon elongatus* and *Aneides hardii*. The arrangement of distal tarsals 4 (d4) and 5 (d5), and the fibulare (f) is diagnostic. c, centrale. Cartilage is stippled; bone and contour of limb are outlined. The pattern observed in *Karsenia* was absent in large samples of *Plethodon* (40 *P. elongatus*, 96 *P. glutinosus*, 98 *P. hoffmani*) and most *Desmognathus* (one asymmetrical variant in *D. monticola*, but absent in 77 others, in 42 *D. quadramaculatus* and 136 *D. ochrophaeus*).



Figure 1 Paratype of *Karsenia koreana* (MVZ 247156) in life.

glossal muscles, lacks posterior flaps. When manipulated in an anaesthetized individual, the tongue formed a well-defined, highly protrusible pad. All teeth are small, undifferentiated and weakly bicuspid. Premaxillary teeth number 4–7 (mean 6.4) in males, 6–10 (mean 7.9) in females; maxillary teeth 43–50 (mean 45.6) in males, 40–56 (mean 49.3) in females; vomerine teeth 13–21 (mean 15.5) in males, 12–20 (mean 17.8) in females.

The species is darkly pigmented with a broad dorsal stripe that varies from being prominent and bright, to being heavily invaded by melanic pigment that obscures the stripe on the head and trunk. The stripe arises as a triangular patch on the snout and extends along the dorsum nearly to the tip of the tail. In life, the stripe is reddish to yellowish brown, tan, or dark brown with some reddish highlights. The stripe is always clearest and brightest in the tail base and anterior tail regions. A canthus rostralis is accentuated by the contrasting colours of the snout. Lateral surfaces are uniformly dark with an obscure overlay of small, whitish speckles. Abundant silvery dots and flecks (about 1 mm in diameter maximum) are concentrated on the dark ventrolateral surfaces and are largest and most numerous where the flanks meet the ventrum. Ventral surfaces are pale, especially in the gular area, and are marked with moderately large, whitish flecks except along the midline.

A single cleared and stained specimen (SIUC H7894) and radiographs of it and seven more (SIUC H7890–7897) reveal that the skull is fully articulated with a small, anteriorly placed intermaxillary foramen and no dorsal fontanelle. Premaxillaries are

paired and relatively stout, with short, expanded frontal processes that do not articulate with each other behind the short foramen. The septomaxillaries are unusually large, articulating on one side with the relatively large prefrontal that lies between but mainly posterior to the nasals and maxillaries. The otic capsule bears small, projection-like lateral crests but no dorsal crests. The operculum has a well-developed, rod-like stylus. The vomers are well formed and bear small, truncated, preorbital processes. Teeth are in short series between the internal nares, usually not on the process. Large, paired patches of paravomerine teeth, well separated from the vomerine teeth, are borne on the ventral surface of the parasphenoid.

Vertebrae include one cervical, 16 ($n = 7$) or 17 ($n = 1$) trunk, one sacral, three caudosacral and a variable number of caudals (maximum 42). Biconcave centra of the trunk vertebrae are relatively stout and posteriorly directed parapophyses are relatively long and slender. The centrum length is shorter than the distance from the tip of one parapophysis to the other. Long bicapitate ribs are about the same length as the centrum and are present on all or all but the last trunk vertebrae.

Limbs are relatively short; a proximal spur is present but partly attached to the shaft of the tibia. Digits in order of decreasing length are 3–2–4–1 and 3–4–2–5–1. Phalangeal formulas are 1–2–3–2 and 1–2–3–3–2. The tarsus and carpus are unmineralized cartilage and cannot be visualized in radiographs. The arrangement of the tarsus in the cleared specimen was confirmed in a dissected specimen (Fig. 2). Tips of terminal phalanges are bluntly rounded, but those of the longest digits can be slightly knobbed.

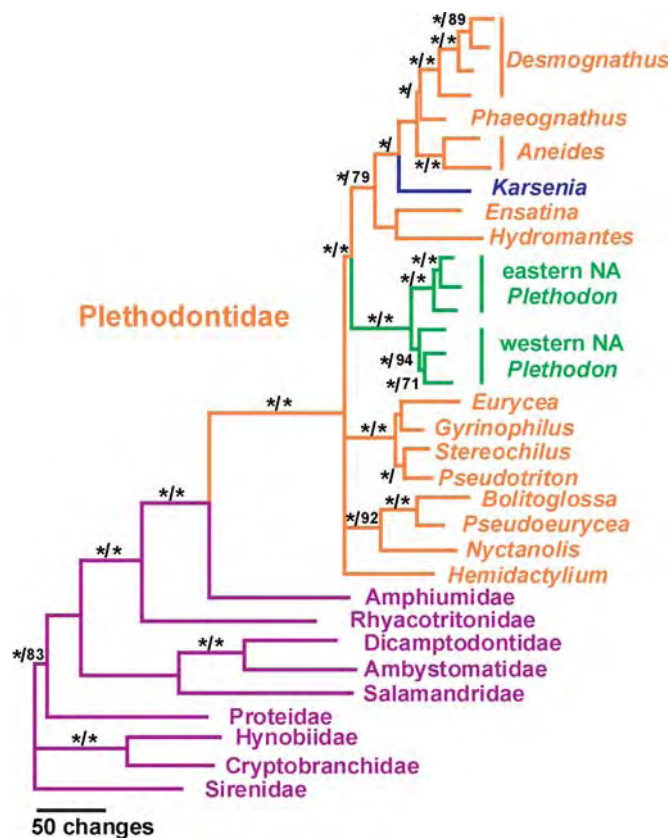
The tongue skeleton is largely cartilaginous and closely resembles that in *Plethodon*, *Aneides* and *Desmognathus*^{4,5}. The first ceratobranchial is the longest element ($1.2 \times$ basibranchial length), whereas the epibranchial is slightly shorter than the basibranchial. The basibranchial, which is mostly ossified, has a pronounced anterior projection in front of the very elongate and slender cornua (which are about $0.6 \times$ basibranchial length). A small, bowed urohyal is ossified (about $0.35 \times$ basibranchial length).

Teeth are moderate in number. In the cleared and stained individual there are four teeth on each premaxilla, 20 and 21 on the two maxillae, 27 on each mandible, five on each vomer, and 70 and 77 in the two well-separated paravomerine patches.

In external morphology *Karsenia* most closely resembles species of *Plethodon* from western North America, but is smaller (closest in size to *P. larselli*, which reaches 57 mm svl⁹). It is shorter-bodied, less robust and has less prominent jaw muscles than members of the *P. elongatus* group from northwestern California and southwestern Oregon, but its hands and feet are nearly identical in shape, length of digits and degree of webbing (Fig. 2). *Karsenia* resembles western *Plethodon* in most features of osteology, including paired premaxillae and general pattern of the skull roof, tongue skeleton and vertebral dimensions. However, its tarsus and vomer resemble those of *Aneides*, from which it differs in critical features of osteology and in having much shorter toes with little or no terminal expansion. Its jaw muscles are much less swollen than in *Aneides*.

Analysis of the nuclear encoded gene *Rag-1* from an array of taxa representing all salamander families shows that *Karsenia* is a plethodontid (Fig. 3). Within a monophyletic Plethodontidae *Karsenia* is nested in a well-supported clade that includes the desmognathine salamanders and *Aneides*. *Hydromantes* and *Ensatina* also have strong affinities to this clade. The sister taxon of all of the above is *Plethodon*, which is monophyletic in agreement with recent molecular studies^{10,11}. There is no support for *Karsenia* as a member of *Plethodon*. *Rag-1* is an evolutionarily conserved gene, and the minimal divergence (4.6% to *Phaeognathus*) is relatively high, greater than that between several other plethodontid genera¹⁰.

The new species occurs in damp, mossy talus slopes and rock-slides of limestone in forests of hardwoods and mixed hardwood/pine (*Pinus densiflora* and *Quercus mongolica* dominate). Forests range from 15 to 50 years in age. Most individuals were found under



small rocks or rock flakes scattered among larger boulders on fine-grained soil.

The discovery of *Karsenia* has important biogeographical implications. Plethodontidae is now seen to have a Holarctic distribution, although it is impressively disjunct. The nearest plethodontids are found along the central coast of British Columbia, Canada (*P. vehiculum*, *Ensatina eschscholtzii*). To the west, no plethodontids are known until central Italy. Disjunctions between Asia and North America are recorded for many taxa, and for plants of eastern Asia and eastern North America. About 65 flowering plant genera¹² display such disjunctions, typically with Asian species in a given clade being more numerous¹³, a strong contrast with plethodontids, which are represented by many species in eastern North America. Phylogenetic relationships of the plants and estimated divergence times suggest multiple historical events at different periods in geological time, with disjunctions typically dating from the Miocene or earlier^{12–14}. Disjunct taxa usually are not cladistically close¹⁴. These conclusions apply to *Karsenia* as well. The species has no close relatives in North America, and the amount of genetic divergence suggests a long period of independent evolution, possibly pre-Tertiary^{8,10}. Plethodontids are known to be very old, with fossils of *Plethodon* and *Aneides* known from the Arikarean (Oligocene) of Montana^{15,16}.

Mammalian fossils indicate strong physical and phylogenetic connections between east Asia and North America at least until the Late Miocene¹⁷, and polar sea temperatures are estimated to have averaged about 15 °C 70 million years ago and more than 20 °C at 90 million years ago¹⁸. Such temperatures would have been favourable for plethodontid salamanders. Accordingly, there have been many opportunities for early migrations between the continents. An important implication of current distributions is that Old World plethodontids have had lower rates of speciation than those in the New World, but also may have been disproportionately subject to extinction. The discovery of plethodontids in Asia encourages further efforts to find more of these secretive animals. □

Received 5 November 2004; accepted 17 February 2005; doi:10.1038/nature03474.

1. *AmphibiaWeb* [online] (<http://amphibiaweb.org>) (Berkeley, California, 2005).
2. Wake, D. B. Comparative osteology and evolution of the lungless salamanders, family Plethodontidae. *Memoirs, South. Calif. Acad. Sci.* **4**, 1–111 (1966).
3. Wake, D. B. & Elias, P. New genera and a new species of Central American salamanders, with a review of the tropical genera (Amphibia, Caudata, Plethodontidae). *Contrib. Sci. Nat. Hist. Mus. Los Angeles Co.* **345**, 1–19 (1983).
4. Lombard, R. E. & Wake, D. B. Tongue evolution in the lungless salamanders, family Plethodontidae. II. Function and evolutionary diversity. *J. Morphol.* **153**, 39–80 (1977).
5. Wake, D. B. & Deban, S. M. in *Feeding* (ed. Schwenk, K.) 95–116 (Academic, San Diego, 2000).
6. Shubin, N. H. & Wake, D. B. in *Amphibian Biology* Vol. 5 (eds Heatwole, H. & Davies, M.) 1782–1808 (Surrey Beatty & Sons, Chipping Norton, Australia, 2003).
7. Wake, D. B., Maxson, L. R. & Wurst, G. Z. Genetic differentiation, albumin evolution, and their biogeographic implications in plethodontid salamanders of California and southern Europe. *Evolution* **32**, 529–539 (1978).
8. Larson, A., Weisrock, D. W. & Kozak, K. H. in *Reproductive Biology and Phylogeny of the Urodela* (ed. Sever, D.) 31–108 (Science Publishers, Enfield, New Hampshire, 2003).
9. Brodie, E. D. Jr Western salamanders of the genus *Plethodon*: systematics and geographic variation. *Herpetologica* **26**, 468–516 (1970).
10. Chippindale, P. T., Bonett, R. M., Baldwin, A. S. & Wiens, J. J. Phylogenetic evidence for a major reversal of life history evolution in plethodontid salamanders. *Evolution* **58**, 2809–2822 (2004).
11. Mueller, R. L., Macey, J. R., Jaekel, M., Wake, D. B. & Boore, J. L. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. *Proc. Natl Acad. Sci. USA* **101**, 13820–13825 (2004).
12. Wen, J. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annu. Rev. Ecol. Syst.* **30**, 421–455 (1999).
13. Xiang, Q.-Y. et al. Regional differences in rates of plant speciation and molecular evolution: a comparison between eastern Asia and eastern North America. *Evolution* **58**, 2175–2184 (2004).
14. Xiang, Q.-Y., Soltis, D. E. & Soltis, P. S. The eastern Asian and eastern and western North American floristic disjunction: congruent phylogenetic patterns in seven diverse genera. *Mol. Phylogenet. Evol.* **10**, 178–190 (1998).
15. Tihen, J. A. & Wake, D. B. Vertebrae of plethodontid salamanders from the Lower Miocene of Montana. *J. Herpetol.* **15**, 35–40 (1981).
16. Rasmussen, D. L. & Prothero, D. R. in *Cenozoic Systems of the Rocky Mountain Region* (eds Raynolds, R. G. & Flores, R. M.) 479–499 (Rocky Mountain Section, Society for Sedimentary Geology, Denver, 2003).
17. Wallace, S. C. & Wang, X. Two new carnivores from an unusual late Tertiary forest biota in eastern North America. *Nature* **431**, 556–559 (2004).

18. Jenkyns, H. C., Forster, A., Schouten, S. & Sinninghe Damsté, J. S. High temperatures in the Late Cretaceous Arctic Ocean. *Nature* **432**, 888–892 (2004).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank S. Karsen for bringing these salamanders to our attention. R. Diaz, K. Klitz and W. Korff assisted with illustrations. Radiographs were obtained with the assistance of B. M. Burr and M. R. Thomas. J. Lazell, R. Mueller, D. Rasmussen, E. Rosenblum, M. Stöck, V. Vredenburg and M. Wake offered assistance and information. Research was supported by a NSF AmphibiaTree grant.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.B.W. (wakelab@uclink.berkeley.edu).

An extant cichlid fish radiation emerged in an extinct Pleistocene lake

Domino A. Joyce^{1,2,3}, David H. Lunt¹, Roger Bills⁴, George F. Turner¹, Cyprian Katongo⁵, Nina Duftner⁵, Christian Sturmbauer⁵ & Ole Seehausen^{1,2,3}

¹School of Biological Sciences, The University of Hull, Cottingham Road, Hull HU6 7RX, UK

²Institute of Zoology- Aquatic Ecology & Macroevolution, University of Bern, Baltzerstr. 6, CH-3012 Bern, Switzerland

³EAWAG, Fish Ecology and Evolution, Limnological Research Center, CH-6047 Kastanienbaum, Switzerland

⁴South African Institute for Aquatic Biodiversity, Private Bag 1015, Grahamstown 6140, South Africa

⁵Department of Zoology, Karl-Franzens-Universität Graz, Universitätsplatz 2, A-8010 Graz, Austria

The haplochromine cichlid fish of the East African Great Lakes represent some of the fastest and most species-rich adaptive radiations known¹, but rivers in most of Africa accommodate only a few morphologically similar species of haplochromine cichlid fish. This has been explained by the wealth of ecological opportunity in large lakes compared with rivers. It is therefore surprising that the rivers of southern Africa harbour many, ecologically diverse haplochromines. Here we present genetic, morphological and biogeographical evidence suggesting that these riverine cichlids are products of a recent adaptive radiation in a large lake that dried up in the Holocene. Haplochromine species richness peaks steeply in an area for which geological data reveal the historical existence of Lake palaeo-Makgadikgadi^{2,3}. The centre of this extinct lake is now a saltpan north of the Kalahari Desert, but it once hosted a rapidly evolving fish species radiation, comparable in morphological diversity to that in the extant African Great Lakes. Importantly, this lake seeded all major river systems of southern Africa with ecologically diverse cichlids. This discovery reveals how local evolutionary processes operating during a short window of ecological opportunity can have a major and lasting effect on biodiversity on a continental scale.

Lake Victoria (LV) and Lake Malawi (LM) are known for their extremely diverse flocks of haplochromine cichlid fish and have received much attention as model systems for explosive speciation and ecomorphological diversification^{1,4}. Together comprising more than 800 species, geological and mitochondrial (mt)DNA evidence suggests that these flocks evolved in a very short time period (Lake Victoria, 15,000–250,000 years^{5,6}; Lake Malawi, up to 5 million years⁷) from one or a few common ancestors^{8,9}. In contrast, the

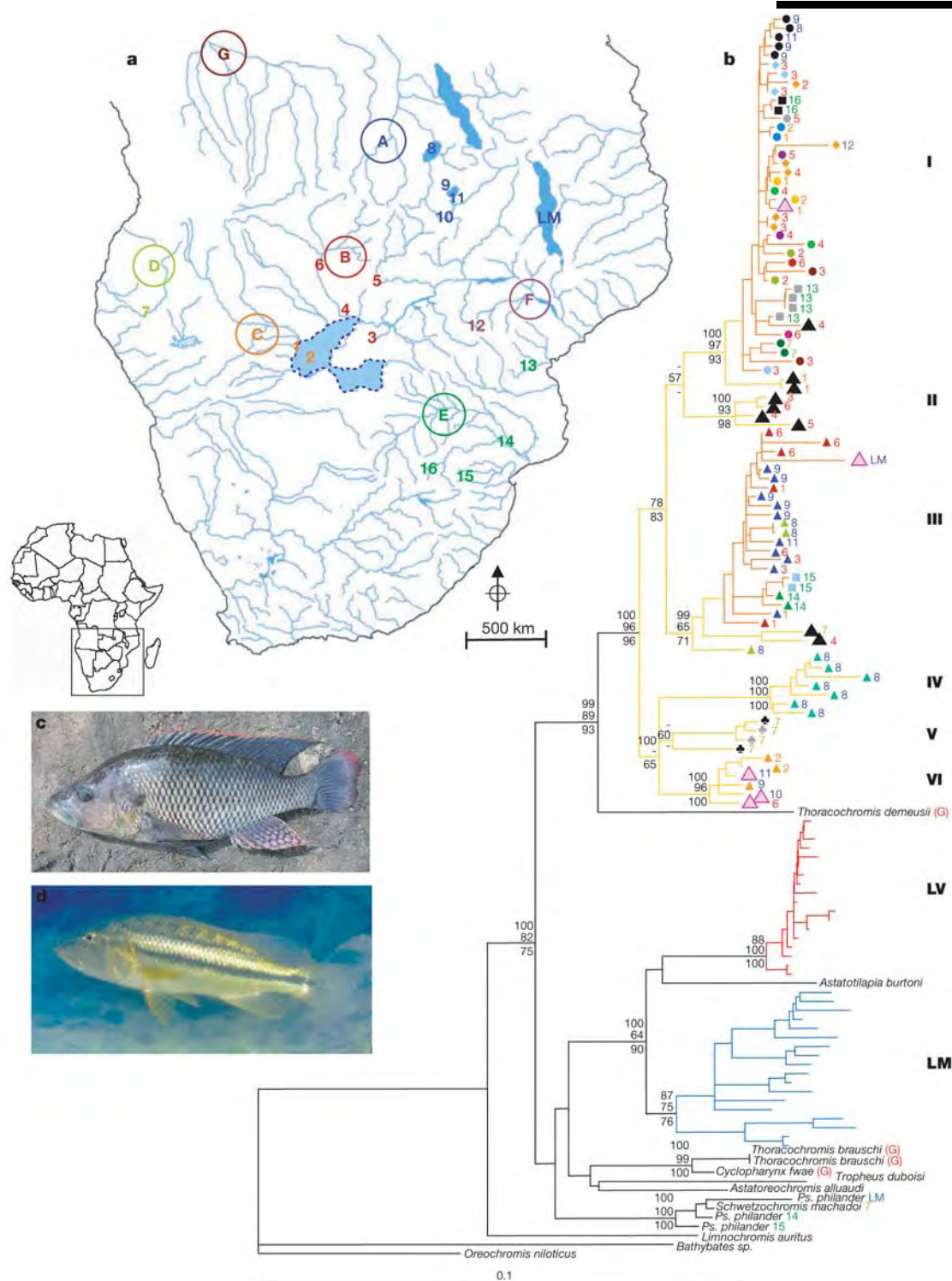


Figure 1 The serranochromine cichlids of southern Africa are a flock of closely related species that probably arose in Lake paleo-Makgadikgadi. **a**, The position and size of Lake palaeo-Makgadikgadi at its Quaternary maximum. A, upper Congo (contains six serranochromine species, one *Thoracochromis* species); B, middle/upper Zambezi (12 serranochromine species); C, Okavango River system (11 serranochromine species); D, Cunene River (seven serranochromine species, two *Thoracochromis* species); E, Limpopo River (four serranochromine species); F, lower Zambezi including Lake Chivero (three serranochromine species); G, middle Kasai to lower Congo (no serranochromine species, but seven *Thoracochromis*-related species, of which four are endemic to Lake Fwa in the middle Kasai River). Numbers indicate sampling sites, and species are listed in the Supplementary Information. A dashed line represents the area covered by lacustrine and fluvio-lacustrine landforms representing Lake Palaeo-Makgadikgadi at its Quaternary maximum. **b**, Neighbour-joining (NJ) tree (GTR + I + Γ) with bayesian posterior probabilities and NJ bootstrap support above the branches, and bootstrap support for maximum likelihood shown below the branches. Only values >50% are shown. The

southern radiation is shown in orange, with rapidly radiating clades (that is, the Lake palaeo-Makgadikgadi species flock) in darker orange, Lake Victoria (LV) clade in red and Lake Malawi (LM) clade in blue. Species are: *Sargochromis mellandi* (black circles), *Sarg. mortimeri* (grey circle), *Sarg. codringtonii* (blue circles), *Sarg. n. sp. 1* (purple circles), *Sarg. carlottae* (yellow circles), *Sarg. n. sp. 2* (bright green circles), *Sarg. giardi* (olive green circles), *Sarg. cf. mortimeri* (red circle), *Sarg. sp.* (brown circles), *Sarg. cf. giardi* (pink circle), *Sarg. coulteri* (dark green circles), *Sarg. cf. giardi2* (light blue circle); *Serranochromis macrocephalus* (large black triangles), *S. robustus* (large pink triangles), *S. altus* (red triangles), *S. angusticeps* (blue triangles), *S. stappersi* (turquoise triangles), *S. meridianus* (bright green triangles), *S. n. sp. Mfimo* (olive green triangles), *S. thumbergi* (light orange triangles); *Chetia flaviventris* (black squares), *C. brevicauda* (grey squares), *C. brevis* (blue squares); *Pharyngochromis* n. sp. (blue diamonds), *P. acuticeps* (orange diamonds); *Thoracochromis albolabris* (black cloverleaves), *T. buysi* (grey cloverleaves). **c**, **d**, *Serranochromis macrocephalus* (**c**) and *S. robustus* (**d**) appear in three of the six southern radiation clades.

ivers associated with Lake Victoria and Lake Malawi show very low haplochromine species richness and morphological diversity; this is believed to be because rivers lack the wealth of ecological opportunity thought to drive adaptive radiation in lakes^{4,10}. There are two widespread riverine lineages of haplochromine cichlids, *Pseudocrenilabrus* and *Astatotilapia/Thoracochromis*, and rivers in this and many other parts of Africa typically contain one species of each¹¹. Paradoxically, the rivers south and southwest of the East African rift valley are markedly different. The upper Congo, middle/upper Zambezi, Okavango, Cunene and Limpopo Rivers each harbour numerous sympatric haplochromines of diverse shape and size, classified as serranochromines¹² and *Thoracochromis* spp. (Fig. 1a).

We collected the majority of the described serranochromine species from these five river systems, as well as seven undescribed species. We also collected four *Thoracochromis* species (Fig. 1a, Supplementary Table S1 and Supplementary Fig. S1) and two other haplochromine cichlids that occur in the area (*Pseudocrenilabrus philander* and *Schwetochromis machadoi*). To reconstruct the phylogenetic origins of the southern African riverine cichlid diversity, we amplified and sequenced the complete mitochondrial control region from 99 specimens and generated phylogenies using minimum evolution, maximum likelihood and bayesian analysis (Fig. 1b). The mtDNA haplotypes of all serranochromine and the two Cunene River *Thoracochromis* species form a well-supported (>95% bootstrap support, 100% bayesian posterior probability) lineage comprising six clades (I–VI, the ‘southern radiation’) to the exclusion of East African haplochromines, *Pseudocrenilabrus*, *Schwetochromis* and Congolese *Thoracochromis*. A *Thoracochromis* species from the lower Congo is the sister taxon to this radiation.

Within this southern radiation, we recover a pattern typical of the adaptive radiations in the East African Great Lakes: emergence of ecological and phenotypic diversity in a rapidly multiplying lineage, characterized by low levels of neutral genetic divergence¹⁰. There are a large number of closely related haplotypes that have arisen from a small number of more divergent haplotypes (Fig. 1b). Clades I and III contain many different genera and species. Within these clades, haplotypes are not sorted by species, and branch lengths are short (mean pairwise genetic distance: clade I = 0.010, s.d. = 0.006; clade III = 0.019, s.d. = 0.010; Lake Victoria = 0.008, s.d. = 0.004; Lake Malawi = 0.042, s.d. = 0.016). Both patterns characterize the species flock of Lake Victoria and sections of the Lake Malawi species flock. The genera *Chetia* and *Serranochromis* contain divergent haplotypes and appear poly- and paraphyletic, respectively. More strikingly, we find deep divergence within *Serranochromis robustus* and *S. macrocephalus*, and even within a single population of *S. macrocephalus* (population 4; mean pairwise genetic distance = 0.036, s.d. = 0.001, across populations = 0.032, s.d. = 0.012). Conversely, several morphologically different and geographically distant species have very similar haplotypes. For example, one *Chetia brevicauda* from the Mussapa River, Mozambique (population 13) and *Serranochromis macrocephalus* from the upper Zambezi (population 4) differ by only 6 out of 887 base pairs (0.68%, genetic distance = 0.0072). At present, these rivers are separated by the middle Zambezi, the Victoria Falls, the lower Zambezi and about 100 km of Indian Ocean shoreline.

These phylogenetic relationships generate a second paradox: a species flock that emerged rapidly and simultaneously in many geographically distant river systems, in which very similar haplotypes occur thousands of kilometres apart, yet for which strongly differentiated haplotypes are found within single populations (Fig. 1b). We suggest a resolution of both paradoxes, based on geological records and biogeographic patterns. We mapped species richness of the serranochromine cichlids across southern Africa using geographical distribution records. We observed a single centre of serranochromine species richness in the area of the Okavango delta and the upper/middle Zambezi (Fig. 2a). Geological records

reveal the presence of a large lake in the Okavango delta region of northern Botswana until about 2,000 yr BP (Fig. 1a), Lake palaeo-Makgadikgadi³. The lake began to form when uplifting along the north-east trending Linyanti and Chobe faults severed the connection between the upper and middle Zambezi River sometime in the Pleistocene (315,000–460,000 yr BP), and diverted the flow of three large rivers, the Cuando, upper Zambezi and Kafue, into an inland drainage basin. This became Lake palaeo-Makgadikgadi¹³. Radiocarbon dating suggests a minimum age of 52,000 yr BP for the lake, and the area covered by lacustrine land forms in northern Botswana amounts to 60,000 km², an area larger than Switzerland.

We suggest that the riverine serranochromine species flock could have arisen in this palaeolake. Climate fluctuations throughout the Pleistocene led to alternating high and low stands, the duration and extent of which are still somewhat uncertain. It seems possible that the lake was largely dry during the last glacial maximum ~18,000 yr BP, when the deep rift Lakes Malawi and Tanganyika experienced 500–600 m lowering of their water levels. At its subsequent Holocene maximum, Lake palaeo-Makgadikgadi was >50 m deep and its potential components (areas of fluviolacustrine surface) amount to a vast 120,000 km² (ref. 14). During the course of the Holocene, however, the lake broke through the fault scarp to the northeast and began to empty into the middle Zambezi, eventually causing the headwaters of the Cuando, upper Zambezi and Kafue to become recaptured by the middle Zambezi. By 2,000 yr BP, the lake had become the seasonal lagoons of the Okavango delta, and further south, the Makgadikgadi saltpans.

To test the hypothesis that the riverine species flock arose in the lost palaeolake against the possibility that this region holds high

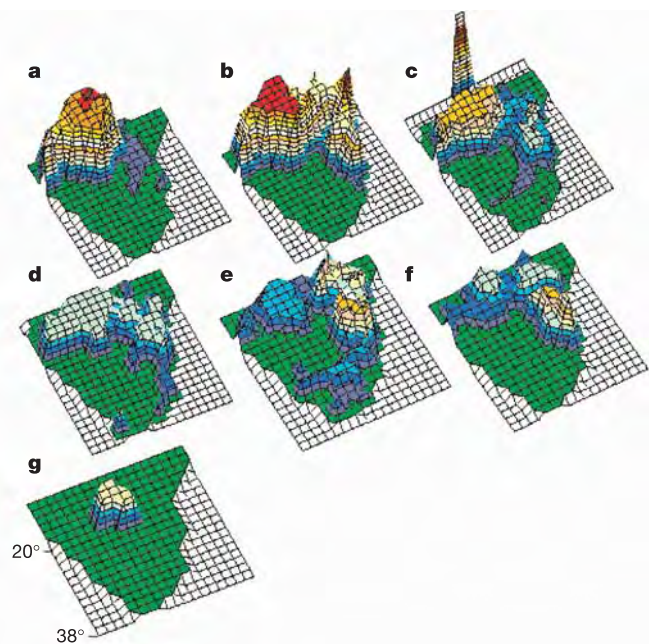


Figure 2 Geographic distribution of species richness in southern African fish groups. x and y axes represent grid squares overlaid on a map of southern Africa, in which occurrence of each species was marked for presence or absence (50% majority rule). The height shows the percentage species of that group found per grid square; gradations are 5% and range from green (0%) dark blue (5–10%) to white (40–45%) to dark red (80–85%). **a**, Southern radiation ($n = 16$ species that are included in our phylogenetic analysis and published distribution maps). **b**, Characidae ($n = 6$). **c**, *Barbus*, the most species-rich African fish group ($n = 43$). **d**, *Oreochromis*, the other main cichlid lineage in southern Africa ($n = 8$). **e**, **f**, Two characteristically riverine fish groups, *Labeo* (**e**, $n = 12$) and *Chiloganis* (**f**, $n = 8$). A single peak of species richness in the Lake palaeo-Makgadikgadi region (**g**) can be seen only for the haplochromines; *Barbus* show a latitudinal gradient which increases into the Congo, represented by a single bar.

riverine fish diversity because of its biogeography and environmental characteristics, we mapped species richness of other fish groups, including three that, like the serranochromines, inhabit river, swamp and lake habitats (the family Characidae, the cyprinid genus *Barbus* and the cichlid genus *Oreochromis*, Fig. 2b–d) and two characteristically riverine groups (the cyprinid genus *Labeo* and the catfish genus *Chiloglanis*, Fig. 2e and f). No other taxon showed a single centre of species richness in the Okavango delta and upper/middle Zambezi, although many rivers now exist there.

Despite the lack of detailed distribution data within the Congo, it is clear that the genus *Barbus* (the most species-rich of all African riverine fish groups) does not peak in the Okavango/Zambezi, but reveals a latitudinal gradient of species richness, in which the percentage of African *Barbus* species increases northwards towards the Equator, from 9% in the Okavango and middle/upper Zambezi to 20% in the upper/middle Congo and even higher in the lower Congo. A similar pattern can be seen in the Characidae. In contrast, serranochromine richness decreases northwards towards the Equator: 43% are in the Okavango and middle/upper Zambezi area, 28% in the upper Congo, 20% in the upper Kasai (Congo system) and none in the lower Congo. Distribution patterns in all of these unrelated fish groups imply connections in the recent past between the Limpopo, middle Zambezi, Okavango, Cunene and upper Congo—the river systems that harbour the remnants of the Lake palaeo-Makgadikgadi cichlid radiation. However, only the species-richness landscape of the serranochromines implies dispersal from a single centre of diversification, the position of which coincides with that of Lake palaeo-Makgadikgadi.

To test the prediction that rapid diversification was driven by the ecological opportunity afforded by new niches associated with the formation of a large lake, we compared ecomorphological variation present in the serranochromine radiation with that in other riverine haplochromines and the Lake Victoria and Lake Malawi radiations. We measured ten linear morphometric distances that reflect ecologically relevant shape variation¹⁵ from between one and ten mature individuals from each species and also from representatives of all major trophic types of Lake Victoria and Lake Malawi cichlids. Principal component analysis was used to extract major axes of variation in the southern radiation, and shows that morphological diversity in the Lake palaeo-Makgadikgadi radiation is directly comparable to that in the haplochromine cichlid radiations of Lake Victoria and Lake Malawi, and is much larger than the variation in East African riverine haplochromines (Fig. 3a).

Ecological feeding types of all three lakes occupy similar positions in morphospace (Fig. 3b). For example, ambush piscivores from Lakes palaeo-Makgadikgadi, Victoria and Malawi all have relatively long heads and long, narrow jaws. Conversely, snail-crushing species have wider, deeper heads, which help to accommodate the hypertrophied pharyngeal jaw bones and muscles required for snail crushing. Absent from the Lake palaeo-Makgadikgadi morphospace are ecotypes that occupy niches now absent from the riverine habitats of the Lake palaeo-Makgadikgadi region: fish with relatively wide heads, shallow bodies and large eyes (represented by the pelagic zooplanktivores in Lake Malawi and Lake Victoria), the most shallow-bodied, long-headed fish (pelagic piscivores of Lake Malawi) and fish with short, wide heads and jaws, which tend to be epilithos-feeding rocky shore dwellers. This is consistent with the hypothesis that these ecological types might not have persisted when their habitats disappeared once the lake dried up.

In Fig. 3, we show the putative ancestors of each flock: the East African riverine haplochromines, some of which must have seeded Lake Malawi and Lake Victoria, and the basal members of the Lake palaeo-Makgadikgadi radiation, including those that possess haplotypes from more than one southern radiation clade (as these species might have originated before the recent radiation). Importantly, the Lake palaeo-Makgadikgadi flock differs from the Lake Malawi and Lake Victoria radiations in historical contingency

because its founders were likely to have been piscivores, rather than fish of a more generalist morphology that would feed on invertebrates and fish fry. Ecological types seem to have evolved in parallel directions along the axes of variation from different ancestral starting points. The ecologically relevant shape variation that accompanied the radiation in Lake palaeo-Makgadikgadi is comparable in magnitude to that of the classic examples of cichlid adaptive radiations⁴ (see Supplementary Fig. S1).

The genetic divergence that we observe between the most divergent mtDNA haplotypes within the southern radiation is comparable to that between the oldest lineages of Lake Malawi cichlids, estimated at a minimum of 1–2 Myr (ref. 16) but possibly up to 5 Myr old. Hence, these mutations cannot have accumulated in the ~50,000–400,000 yr since Lake palaeo-Makgadikgadi formed, but must have existed in the ancestral population. In order to retain the deep polymorphism observed (for example, across *S. robustus* and *S. macrocephalus*), this population would have had to maintain a substantial size, not only throughout the Lake palaeo-Makgadikgadi colonization and diversification period, and while the lake fluctuated in size, but throughout the much longer time between the occurrence of the deepest split in the southern

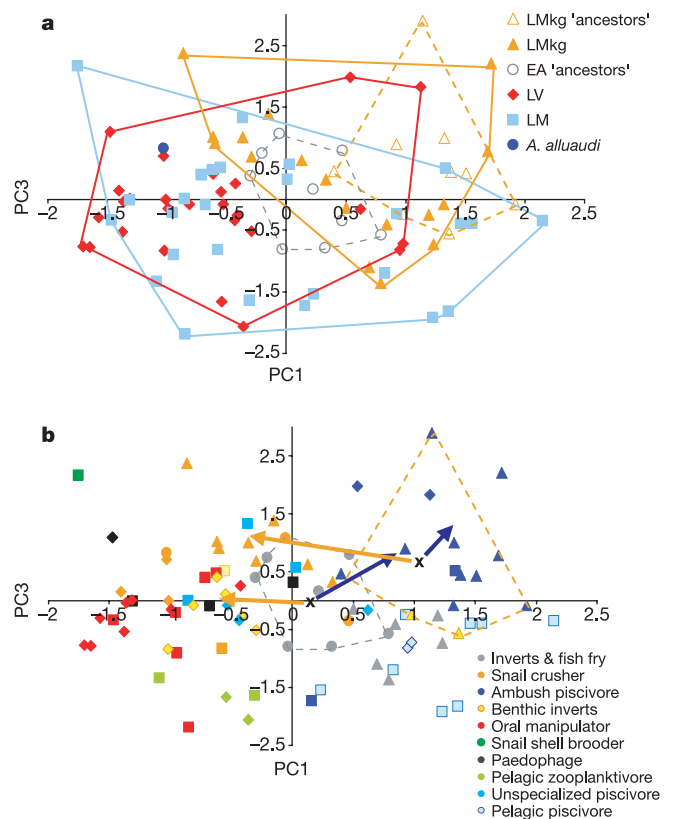


Figure 3 The area in morphospace occupied by the cichlid radiations. **a**, Most differentiation between the river and lake groups is along principal components (PC) 1 and 3. PC1 increases with decreasing interorbital width, increasing head length and lower jaw length-to-width ratio. PC3 increases with an increase in cheek depth. The genetically diverse ancestral taxa (east African (EA) rivers and species of the southern radiation that contain the oldest haplotypes) occupy a smaller volume in morphospace than the genetically similar taxa that only emerged in the recent radiations. Species from Lake palaeo-Makgadikgadi (LMkg) occupy a volume comparable to species from Lake Victoria (LV) and Lake Malawi (LM). **b**, Colours represent trophic groups; symbols represent lakes (EA 'ancestors', circles; LMkg, triangles; LV, diamonds; LM, squares). Fish of similar feeding types from different lakes have diverged from their ancestors in similar directions, as expected if the radiation was driven by divergent ecological adaptation (snail-crushers and ambush piscivores are shown with appropriately coloured arrows).

haplotype clade and the formation of Lake palaeo-Makgadikgadi. Alternatively, the lake (which connected many drainage systems¹⁷) could have provided an opportunity for secondary contact, such that allopatrically divergent lineages of *Serranochromis* could meet and hybridize at the onset of the radiation¹⁸. The products of the radiation could then have dispersed through the Late Pleistocene and Holocene drainage systems to generate the present geographical distribution patterns of haplotype diversity. *In situ* radiation from several divergent ancestral haplotypes, followed by dispersal out of the radiation site, offers an explanation for why some geographically distant and morphologically divergent species have more similar haplotypes than some conspecific individuals from the same locality.

We conclude that the southern African riverine cichlids represent a case of explosive adaptive radiation into a functionally diverse species flock that emerged rapidly in a short, geologically transient window of opportunity. During the unstable history of Lake palaeo-Makgadikgadi and its drainage systems¹⁷, the members of this flock spread through the southern African rivers, where a substantial number of them have persisted and are now the only remnants of the radiation after the lake dried up ~2,000 yr BP. The origins of the functional diversity we now observe may have needed the ecological opportunities provided by a large lake, but the south African rivers seem to have maintained a significant portion of this diversity subsequently. Therefore, this appears to be an example of how a localized evolutionary process—an adaptive radiation in a geologically transient environment—can have a profound and lasting effect on the ecological and genetic diversity of a continental fauna. □

Methods

Taxon sampling

Sample taxa and site data can be found in Supplementary Table S1. We collected 17 out of 26 described *serranochromines*, and seven as yet undescribed *serranochromine* species from five river systems (Fig. 1a), as well as 4 out of the 8 *Thoracochromis* species of the Congo and Cunene, collecting the same species from more than one of these major river systems where they occur in multiple systems. We also collected the only two other haplochromine cichlids of southern Africa (*Pseudocrenilabrus philander* and *Schwetochromis machadoi*), as well as *Astatotilapia calliptera* from Lake Malawi. We used representatives of Lake Malawi and Lake Victoria haplochromines and other outgroups from GenBank (see Supplementary Table S1). Whole specimens of most of the southern African populations, except those from the Congo River, are held in the collection of SAIAB (South African Institute of Aquatic Biodiversity).

DNA extraction, amplification and sequencing

DNA was extracted from fin clips or muscle tissue using a standard phenol-chloroform extraction procedure, and the entire control region was amplified using forward primer HapThr-2 + 4 (5'-CCTACTCCCAAGCTAGGATC-3') and reverse primer Fish12s, (5'-TGCCGAGACTTGTCATGTGTAAG-3'). Polymerase chain reaction (PCR) products were cleaned using exonuclease and shrimp alkaline phosphatase, and a combination of the amplification primers and two internal primers (forward primer Dloopint 5'-AGCCAC CATCAGTTGATT-3' and reverse primer HapDloop 5'-GGTTGCAGG AGTCTTAGAG-3') were used in cycle sequencing. This was carried out using DTCS quickstart (Beckman Coulter) according to the manufacturer's instructions, adding 1 M betaine to the sequencing reaction. Sequences were resolved using a Beckman capillary sequencer.

Sequence alignment and phylogenetic analysis

ClustalX¹⁹ was used to align the sequences (887 bp), with subsequent adjustments made by eye. We have submitted the alignment to GenBank (AY913844–AY913942). We used Modeltest²⁰ to estimate the likelihood parameters for our data set (general time reversible (GTR) substitution model with among site rate heterogeneity following a gamma plus invariant sites distribution, GTR + I + Γ , shape parameter 0.6799) and PAUP* (v4.b10, ref. 21) to build a neighbour-joining tree with 1,000 bootstrap pseudoreplicates. We obtained a tree and posterior probabilities for nodes with MrBayes²² using 2,000,000 Markov chain Monte Carlo generations. We plotted likelihood values against tree number to show that discarding the first 10% of trees ensured stationarity had been reached. A maximum likelihood tree was constructed using PHYML²³ with parameters from Modeltest; 100 pseudoreplicates gave bootstrap values for nodes.

Morphometric analysis

Standard length (SL), body depth (BD), head length (HL), snout length (SnL), pre-orbital depth (PoD), eye length (EyL), cheek depth (CD), inter-orbital width (IoW), lower jaw length (LJL) and lower jaw width (LJW) were measured from a total of 99 *serranochromine* specimens, comprising 23 species from 32 populations (mean of three fish per population), and also from 16 specimens (6 species) of East and North African riverine haplochromines, 6 specimens from one Lake Victoria species, and 15 species of

Lake Malawi cichlids. Data for 26 other species from Lake Victoria, 13 from Lake Malawi, and one from East African rivers were taken from other sources^{24–29}. BD and HL were standardized with SL. SnL, PoD, EyL, and IoW were standardized with HL. The LJL/LJW ratio was used to enable the use of published data for Lake Victoria and Lake Malawi cichlids. HL/SL and LJL/LJW did not show any allometric trend. To remove allometry from the variation in the other distances, we regressed them against SL and used the residuals for further analysis. Because of significant deviation from normality, we log-transformed residual PoD/HL and LJL/LJW. Conservatively, missing data were replaced with means. Principal components 1 and 3 were plotted in Fig. 3.

Biogeographical analysis

We used published distribution data³⁰ for the species of southern Africa and overlaid grid rectangles (150 × 110 km), in which we recorded the presence or absence of each species using a 50% presence majority rule. We calculated the percentage species of each group that occurs per square. We plotted these values against longitude and latitude, assigning marine grids a minus number for clarity. Detailed published distribution data ends at the boundary between the upper Zambezi and the upper Congo, so to elucidate the species richness trends for *serranochromines* and barbs northwards of the Lake palaeo-Makgadikgadi region, we used records of each species from FishBase (<http://www.fishbase.org>) to calculate the percentage of species that occur in the Congo River, represented by a single bar for the barbs and hidden behind the richness peak for the *serranochromines* in Fig. 2.

Received 27 December 2004; accepted 22 February 2005; doi:10.1038/nature03489.

- Kocher, T. D. Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Rev. Genet.* **5**, 288–298 (2004).
- Thomas, D. S. G. & Shaw, P. A. Late Quaternary environmental change in central southern Africa: new data, synthesis, issues and prospects. *Quaternary Sci. Rev.* **21**, 783–797 (2002).
- Thomas, D. S. G. & Shaw, P. A. *The Kalahari Environment* (Cambridge Univ. Press, Cambridge, 1991).
- Fryer, G. & Iles, T. D. *The Cichlid Fishes of the Great Lakes of Africa. Their Biology and Evolution* (Oliver and Boyd, Edinburgh, 1972).
- Johnson, T. C. et al. Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* **273**, 1091–1093 (1996).
- Fryer, G. Biological implications of a suggested Late Pleistocene desiccation of Lake Victoria. *Hydrobiologia* **354**, 177–182 (1997).
- Ebinger, C., Deino, A., Drake, R. & Thesa, A. Chronology of volcanism and rift basin propagation: Rungwe volcanic provinces, East Africa. *J. Geophys. Res.* **94**, 15783–15803 (1989).
- Nagl, S. et al. The origin and age of haplochromine fishes in Lake Victoria, East Africa. *Proc. R. Soc. Lond. B* **267**, 1049–1061 (2000).
- Meyer, A., Kocher, T. D., Basasibwaki, P. & Wilson, A. C. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial-DNA Sequences. *Nature* **347**, 550–553 (1990).
- Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, Oxford, 2000).
- Katongo, C., Kobl Müller, S., Dufter, N., Makasa, L. & Sturmbauer, C. Phylogeography and speciation in the *Pseudocrenilabrus philander* species complex in Zambian Rivers. *Hydrobiologia* (in the press).
- Greenwood, P. A review of the *Serranochromine* cichlid fish genera, *Pharyngochromis*, *Sargochromis*, *Serranochromis* and *Chetia*. *Bull. Nat. Hist. Mus. Zool. Ser.* **59**, 33–44 (1993).
- Moore, A. E. & Larkin, P. Drainage evolution in south-central Africa since the breakup of Gondwana. *S. Afr. J. Geol.* **104**, 47–68 (2001).
- Rognon, P. Pluvial and arid phases in the Sahara: the role of non-climatic factors. *Paleoecol. Afr.* **12**, 45–62 (1980).
- Barel, C. D. N., van Oijen, M. J. P., Witte, F. & Witte-Maas, E. An introduction to the taxonomy and morphology of the haplochromine cichlidae from Lake Victoria. *Neth. J. Zool.* **27**, 381–389 (1977).
- Sturmbauer, C., Baric, S., Salzburger, W., Rüber, L. & Verheyen, E. Lake level fluctuations synchronize genetic divergence of cichlid fishes in African lakes. *Mol. Biol. Evol.* **18**, 144–154 (2001).
- Shaw, P. After the flood: The fluvio-lacustrine landforms of northern Botswana. *Earth Sci. Rev.* **25**, 449–456 (1988).
- Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207 (2004).
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882 (1997).
- Posada, D. & Crandall, K. A. Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818 (1998).
- Swofford, D. L. *PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods)* (Sinauer Associates, Sunderland, Massachusetts, 2003).
- Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
- Guindon, S. & Gascuel, O. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
- Seehausen, O., Lippitsch, E., Bouton, N. & Zwennes, H. Mbipi, the rock-dwelling cichlids of Lake Victoria: description of three new genera and fifteen new species (Teleostei). *Ichthyological Explor. Freshwat.* **9**, 129–228 (1998).
- Seehausen, O. *Lake Victoria Rock Cichlids—Taxonomy, Ecology and Distribution* (Verduijn Cichlids, Zevenhuizen, Netherlands, 1996).
- Greenwood, P. H. *The Haplochromine Fishes of the African Great Lakes* (Krauss Intl., Munich, 1980).
- Eccles, D. H. & Trewas, E. *Malawian Cichlid Fishes. The Classification of Some Haplochromine Genera* (Lake Fish Movies, Herten, Germany, 1989).
- Turner, G. F. Description of a commercially important pelagic species of the genus *Diplotaxodon* (Pisces: Cichlidae) from Lake Malawi, Africa. *J. Fish Biol.* **44**, 799–807 (1994).
- Roberts, T. R. & Kullander, S. O. Endemic cichlid fishes of the Fwa River, Zaïre: systematics and ecology. *Ichthyological Explor. Freshwat.* **5**, 97–154 (1994).
- Skelton, P. A. *A Complete Guide to the Freshwater Fishes of Southern Africa* (Struik, Cape Town, 2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We acknowledge the financial support of the Leverhulme Trust for this project, and that of the Natural Environmental Research Council for Hull's Genome Analysis Suite. The Africa Museum of Belgium and the OEAD of Austria supported C.K.'s fieldwork in Zambia. N.D. received a DOC-FFORTE (women in research and technology) fellowship from the Austrian Academy of Sciences and was further financially supported by the University of Graz. C.S. was supported by the Austrian Science foundation. We also thank A. Smith, P. Skelton, D. Tweddle and W. Salzburger for fin clips, H. J. Mrosso for morphometric measurements of three Lake Victoria cichlid species, A. Konings for *Serranochromis robustus* in Fig. 1d, and K. Young and U. Schliwien for comments and discussion.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to O.S. (ole.seehausen@aqua.unibe.ch). The alignment and DNA sequences presented in this paper are available at GenBank (accession numbers AY913844–AY913942).

Epistasis and balanced polymorphism influencing complex trait variation

Juergen Kroymann & Thomas Mitchell-Olds

Max Planck Institute for Chemical Ecology, Department of Genetics & Evolution, Hans-Knoell-Str. 8, D-07745 Jena, Germany

Complex traits such as human disease, growth rate, or crop yield are polygenic, or determined by the contributions from numerous genes in a quantitative manner. Although progress has been made in identifying major quantitative trait loci (QTL), experimental constraints have limited our knowledge of small-effect QTL, which may be responsible for a large proportion of trait variation^{1–3}. Here, we identified and dissected a one-centimorgan chromosome interval in *Arabidopsis thaliana* without regard to its effect on growth rate, and examined the signature of historical sequence polymorphism among *Arabidopsis* accessions. We found that the interval contained two growth rate QTL within 210 kilobases. Both QTL showed epistasis; that is, their phenotypic effects depended on the genetic background. This amount of complexity in such a small area suggests a highly polygenic architecture of quantitative variation, much more than previously documented⁴. One QTL was limited to a single gene. The gene in question displayed a nucleotide signature indicative of balancing selection, and its phenotypic effects are reversed depending on genetic background. If this region typifies many complex trait loci, then non-neutral epistatic polymorphism may be an important contributor to genetic variation in complex traits.

Because of experimental constraints, our knowledge of complex trait variation is limited to QTL of relatively large effect. Conventional QTL mapping typically uses markers with an average spacing of several centimorgans, and uses sample sizes insufficient to identify minor QTL. Thus, although many minor QTL contribute to complex trait variation^{3,5}, we have little information on the magnitude, frequency, epistasis and fitness consequences of the genes that matter in quantitative genetics. Genes of large effect may not be representative of the majority of segregating polymorphisms, hence studies are needed that avoid the ascertainment bias that results from a focus on large-effect QTL. For this reason, we fine-mapped phenotypic effects segregating within a one-centimorgan chromosome interval for which lines with mapped recombination breakpoints were available, and examined the sequence signature of historical polymorphism.

Recently, we investigated 58 near-isogenic *Arabidopsis* lines that had recombined in a 210-kb interval on the upper arm of

chromosome 5 (refs 6, 7). These lines were developed from a cross between the *Arabidopsis* accession Col-0 and a Col-0/*Ler-0* recombinant inbred line⁸, CL5. We tested for potential allocation costs associated with variation in herbivore resistance, and found no evidence for a growth rate QTL in the central portion of the interval. In the centre of this interval are located *MAM* genes, which encode methylthioalkylmalate synthases, enzymes involved in the synthesis of plant defence compounds^{6,7}. However, two regions several dozens of kilobases upstream and downstream of the *MAM* cluster displayed statistical support for growth rate QTL (Fig. 1). We investigated whether these regions indeed influence *Arabidopsis* biomass accumulation with a variety of advanced crosses that harboured recombination breakpoints in the vicinity of both presumptive QTL. We selected near-isogenic parental lines such that their *F*₂ progeny segregated in a very small interval but not in the flanking regions. We were thus able to investigate the effect of the segregating interval on growth rate.

We fine-mapped the downstream QTL in two series of crosses (Fig. 2 and Table 1). We examined a 30-kb region with progeny from five crosses (grLC1 to grLC5), each with *Ler-0*-like sequence 5' and Col-0 sequence 3' flanking the segregating interval. We also investigated a 13-kb subset of this region with four crosses (grCL6 to grCL9) in the opposite direction. In addition, we generated a control cross, grCL(6–9), in which the entire 13-kb sequence portion was segregating. Both series of crosses revealed the presence of a QTL. In the grLC crosses, the QTL mapped to a 6.7-kb interval segregating in cross grLC2 (*N* = 273; degrees of freedom (d.f.) = 1, 10; *F* = 6.14; *P* < 0.05). This interval contains two complete genes (*At5g23160* and *At5g23170*) and 5' partial sequence from a third gene (*At5g23180*). In the grCL crosses, the QTL mapped to a 1.3-kb interval segregating in the grCL7 cross (*N* = 391; d.f. = 1, 9; *F* = 6.24; *P* < 0.05), which is contained within the 6.7-kb grLC2 interval. The grCL7 interval contains partial sequence of only one gene (*At5g23170*), indicating that this portion of the *Arabidopsis* genome constitutes the growth rate QTL. Only one-third of the 6.7-kb grLC2 interval is not included in any of the reciprocal grCL intervals. grCL6 and grCL8 bear no indication of a growth rate QTL. The remaining sequence between LJ01C and CJ10L (Fig. 2) harbours only three polymorphisms: a synonymous nucleotide substitution within the coding region of *At5g23170*, a 1-base pair

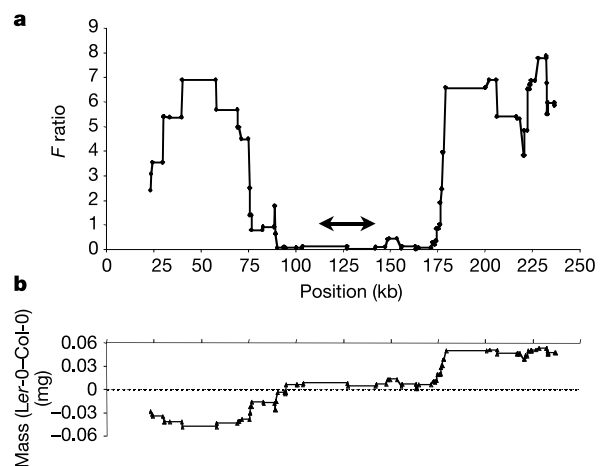


Figure 1 Growth rate differences between Col-0 and *Ler-0* alleles in a 210-kb region of the *Arabidopsis* genome. Position (in kb) is based on an alignment of complete sequences in this region from both accessions. **a**, Statistical significance for the presence of growth rate QTL in near-isogenic lines^{6,7} (see also Supplementary Data and Supplementary Methods). The location of the *MAM* cluster is indicated by an arrow. **b**, Dry weight differences between Col-0 and *Ler-0* genotypes, based on least square means. Note that in this experiment genotype performance is also influenced by uncharacterized growth rate QTL outside the 210-kb interval.

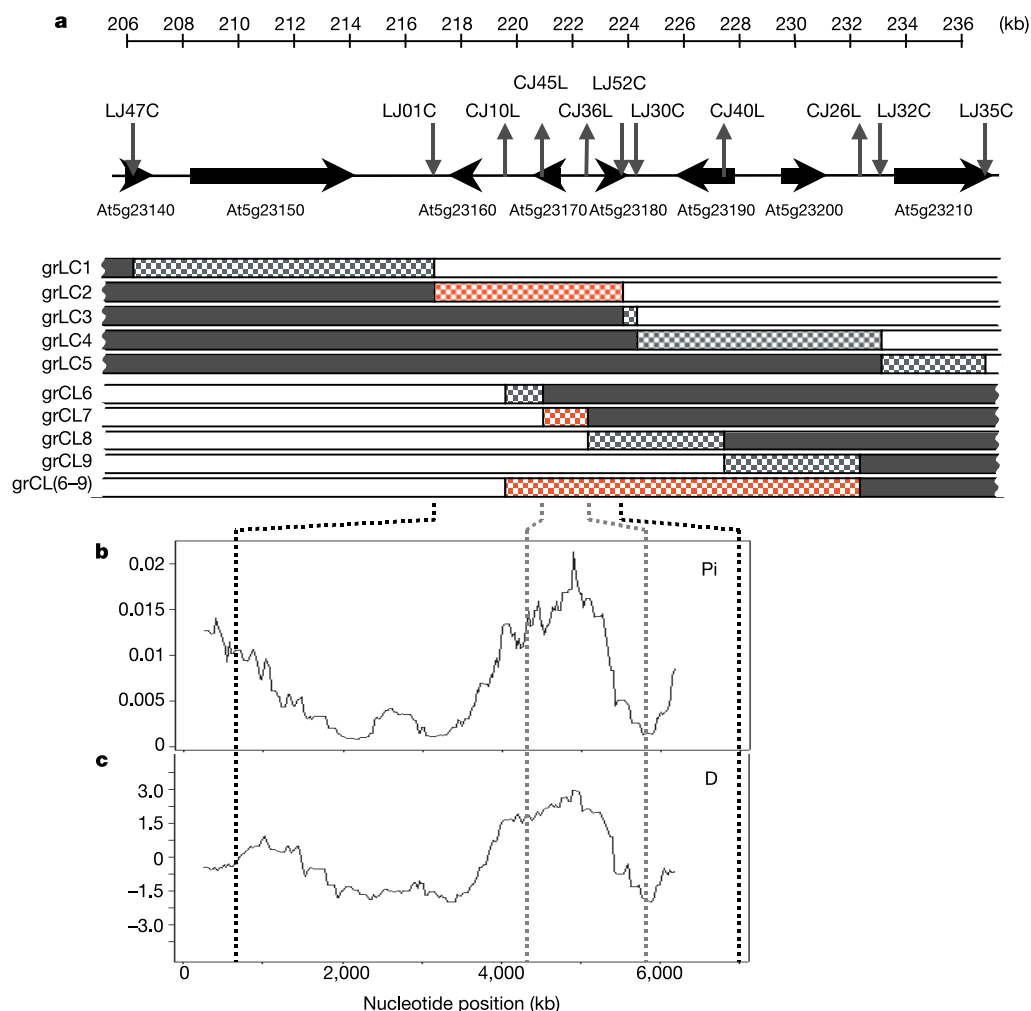


Figure 2 Fine map of the growth rate QTL downstream of the *MAM* cluster. **a**, Horizontal arrows indicate genes (see <http://www.arabidopsis.org>). Vertical arrows indicate recombination breakpoints in the parental lines used for advanced crosses: downwards pointing arrows, *Ler-0* sequence upstream and *Col-0* sequence downstream of the recombination breakpoint; upwards pointing arrows, *Col-0* upstream and *Ler-0* downstream. Intervals segregating (grey and white and red and white checked pattern) in

advanced crosses (grLC1 to grLC5, grCL6 to grCL9 and grCL(6-9)) and flanking genotypes (grey bars, *Ler-0*; white bars, *Col-0*) are shown below. The red and white checked intervals contain the QTL. **b**, **c**, Sliding window analyses of nucleotide polymorphism (**b**, P_i) and Tajima's D (**c**) around At5g23170, based on sequences from 31 *Arabidopsis* accessions. Window size, 500 bp; step width, 25 bp.

(bp) insertion/deletion (indel) in an intron-located adenine mononucleotide stretch, and a single-nucleotide polymorphism (SNP) 324 nucleotides downstream of the At5g23170 stop codon. None of these is likely to cause a QTL. Thus, in conclusion, both the grLC and the grCL crosses identify the same sequence interval as a QTL.

The only gene in this interval, At5g23170, is a putative serine/threonine protein kinase (<http://www.arabidopsis.org>) with substantial similarity to other plant protein kinases^{9,10}. These proteins are usually involved in signal transduction pathways; for example, during cell differentiation and plant development^{9,11} or disease resistance¹². In the PlantsP database (<http://plantsp.sdsc.edu>), At5g23170 is grouped into the receptor-like cytoplasmic kinase family. According to available Macroarray and Affymetrix chip data (<http://plantsp.sdsc.edu>), At5g23170 is transcribed at low intensity throughout the plant, with elevated transcript levels in mature stamens and pollen.

This QTL shows strong epistasis: the direction of allelic effects is reversed depending on genetic background. In the grLC2 cross, the *Ler-0* allele showed higher biomass accumulation, displaying a 15.1% increase in biomass compared with *Col-0*. In contrast, in the grCL7 cross the *Ler-0* genotype was inferior and accumulated on average 5.1% less biomass than the *Col-0* genotype. We confirmed

this result in the control cross grCL(6-9), in which plants with the *Ler-0* genotype performed, on average, 5.5% worse than those with the *Col-0* genotype ($N = 340$; d.f. = 1, 12; $F = 3.74$; $P < 0.05$). Thus, we identified a growth rate QTL within a putative serine/threonine protein kinase whose effects are reversed depending on genetic background.

The upstream QTL (Fig. 3 and Table 1) maps to a region of low recombination. Therefore, we designed crosses to confirm the existence of a QTL, but not locate its exact position. This QTL mapped to a 32-kb interval containing nine genes. Notably, one of these genes, a paralogue of At5g22860 that is annotated as a prolylcarboxypeptidase (E.C. 3.4.16.2) (<http://www.arabidopsis.org>), resides within an 8.5-kb indel polymorphism absent from the *Col-0* accession. We detected the growth rate QTL only when the 5' flanking sequence was *Ler-0*-like and the 3' flanking sequence *Col-0*-like ($N = 314$; d.f. = 1, 14; $F = 9.77$; $P < 0.01$); that is, in the glLC cross, but not vice versa (glCL) ($N = 225$; d.f. = 1, 5; $F = 1.34$; not significant). In the glLC cross, the *Ler-0* allele produced 11.8% more biomass than the *Col-0* allele. Although the two crosses were not exactly reciprocal, only a few polymorphisms segregate in the glLC intervals that are not contained in glCL. Nearly all of these polymorphisms reside in the intergenic region between At5g22890

Table 1 Genotype performance in advanced crosses

Cross	Total (N)*	Col-0 (N)	Ler-0 (N)	LSM (Col-0 fresh weight) (g)	LSM (Ler-0 fresh weight) (g)	Relative fresh weight difference (%)	F ratio	P-value
gILC	639	142	172	0.1156	0.1292	11.8	9.77	0.005 < P < 0.01
gICL	470	101	124	0.1019	0.0969	-4.9	1.34	NS
grLC1	441	93	111	0.1662	0.1675	0.8	0.03	NS
grLC2	527	136	137	0.1396	0.1607	15.1	6.14	0.025 < P < 0.05
GrLC3	611	146	142	0.1613	0.1686	4.5	1.27	NS
grLC4	667	167	156	0.1700	0.1699	-0.1	0.00	NS
grLC5	498	118	145	0.1649	0.1691	2.6	0.22	NS
grCL6	164	39	43	0.1390	0.1307	-6.0	5.97	NS
grCL7	762	196	195	0.1987	0.1885	-5.1	6.24	0.025 < P < 0.05
grCL8	154	30	43	0.1331	0.1340	0.7	3.86	NS
grCL9	834	209	210	0.1983	0.1937	-2.3	0.55	NS
grCL(6-9)	729	175	165	0.1179	0.1114	-5.5	3.74	0.025 < P < 0.05

Relative fresh weight difference was calculated as the difference in fresh weight between Ler-0 and Col-0 genotypes at the segregating interval, divided by the fresh weight of the Col-0 genotype. Numerator mean squares always had 1 degree of freedom; degrees of freedom for the denominator ranged from 1 to 14, with an average of 8. P-values are given for two-tailed tests, except for the control cross grCL(6-9), where a one-tailed test was performed. LSM, least square means; NS, not significant.

*Total for Col-0, Ler-0 and heterozygotes.

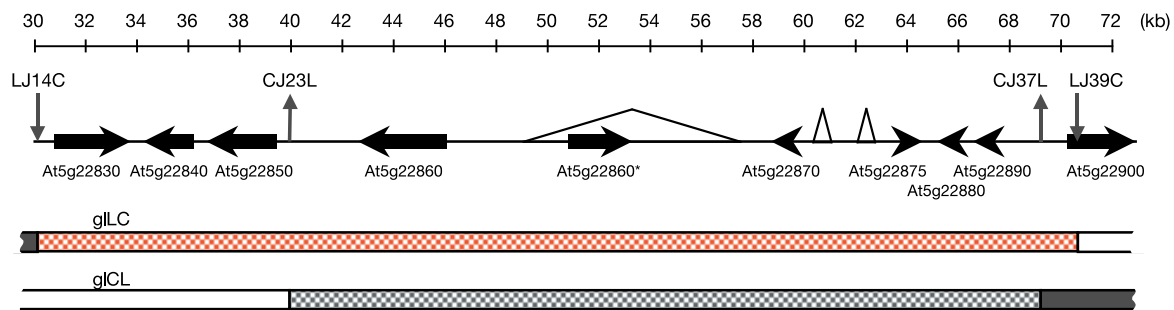


Figure 3 Fine map of the growth rate QTL upstream of the *MAM* cluster. Open triangles indicate major deletions in Col-0 compared to Ler-0. One of these indel polymorphisms

contains an At5g22860 paralogue (indicated by an asterisk). In this region, genotype performance was compared for only two advanced crosses (gILC and gICL).

and At5g22900; that is, in the 3' non-overlapping interval. Only three polymorphisms (two SNPs and a 1-bp indel) are located in any of the three genes encoded in the 5' non-overlapping interval, and none of them alters the open reading frames; all three polymorphisms are intron-located. Therefore, these data suggest that the upstream QTL resides within the 32-kb interval shown in Fig. 3, and shows epistasis with its genetic background.

To estimate genetic background effects, we randomized equal amounts of F₂ seeds from near-isogenic crosses grLC2 and grCL7 and grew them together. On average, grCL7 plants accumulated 34% more biomass than grLC2 plants ($N = 488$; $F = 74.68$; d.f. = 1, 474; $P < 0.0001$), thus showing linked growth rate QTL in the genetic background outside our 210-kb interval. These may correspond to other, previously identified growth-related QTL⁴.

We analysed nucleotide polymorphism patterns at the downstream QTL for an approximately 6.4-kb fragment containing At5g23160 and At5g23170 from 31 *Arabidopsis* accessions (Fig. 2b, c; see also Supplementary Figs 1 and 2). Sliding window analyses revealed high levels of nucleotide polymorphism at and surrounding At5g23170 (Fig. 2b). We found no evidence for an elevated mutation rate when comparing At5g23170 with the closely related *Arabidopsis lyrata*¹³. Instead, elevated nucleotide polymorphism at this QTL represents a signature typical of balancing selection in *A. thaliana*¹⁴. In addition, Tajima's test¹⁵ indicated an excess of intermediate frequency polymorphisms in this region (Fig. 2c). Coalescent simulations confirmed statistical significance for positive values of Tajima's *D* for the At5g23170 coding region, even when applying the stringent criterion of no recombination ($D = 2.17$; $P < 0.05$). This significance test is conservative because genome-wide sequence surveys among *Arabidopsis* accessions typically detect negative values for Tajima's *D*, presumably due to complex demographic processes¹⁶. Our result was also significant in comparison to the genome-wide empirical null distribution of 180 independent *Arabidopsis* loci¹⁶ ($P < 0.02$). Moreover,

At5g23170 maps to a region with high local recombination frequency such that effective selection against deleterious alleles can be expected. Therefore, high levels of nucleotide polymorphism and Tajima's *D* both are confined to a fairly small region and decline sharply in the regions flanking At5g23170.

Initially, we had chosen this 210-kb interval to investigate genes controlling biochemical variation for secondary metabolites^{6,7}, without knowledge or consideration of variation in growth rate. The region has typical values for gene density, G+C content, recombination frequency and levels of linkage disequilibrium^{17,18}. The functional polymorphisms we describe would be undetectable by standard QTL methods: they are too small, they are tightly linked with opposing effects, and the phenotypic effect of the downstream QTL reverses direction owing to epistasis with the genetic background. This QTL is delimited to a single gene that shows the sequence signature of balancing selection. Although the importance of cryptic genetic variation has been recognized in agriculture¹⁹ and human disease²⁰, its genome-wide prevalence has been unclear. If this chromosome interval is representative of most of the genome, then we predict a highly polygenic, epistatic architecture of complex traits, with cryptic genetic variation maintained in part by balancing selection. □

Methods

Crosses

We crossed Col-0 to CL5 (Nottingham *Arabidopsis* Stock Center accession N1901), a recombinant inbred line⁸ derived from a cross between Col-0 and Ler-0. In CL5, approximately 70% of the markers display the Col-0 genotype. Importantly, CL5 does not carry the *erecta* mutation²¹, which is known to influence plant size and morphology. We performed all subsequent crosses with progeny derived from a single F₁ plant from this original cross. We screened F₂ progeny obtained by selfing from the F₁ of this cross for recombination in an approximately 210-kb region flanked by polymorphic microsatellite markers⁵. Initially, we identified 95 recombinants. For 68 recombinants, we mapped recombination breakpoints between their two closest flanking polymorphic sites. As we sequenced the whole region from Ler-0, we knew all polymorphisms distinguishing Col-0

and *Ler-0* in this region. We selected F_4 plants homozygous for the recombination for crosses determined to fine map growth rate QTL such that in each of the resulting F_1 plants a small heterozygotic region was flanked by *Ler-0*-like sequence on one side, and by Col-0-like sequence on the other. As an example, line LJ47C harboured a recombination breakpoint upstream of At5g23150, and has *Ler-0* sequence to the left and Col-0 sequence to the right of this breakpoint (Fig. 2). Line LJ01C had a recombination breakpoint downstream of At5g23150, also with *Ler-0* to the left and Col-0 to the right. We crossed these two lines to generate grLC1. grLC1 F_1 progeny are homozygous *Ler-0* left of the recombination breakpoint in LJ47C and homozygous Col-0 right of the recombination breakpoint in LJ01C, whereas the interval flanked by these breakpoints and containing At5g23150 is heterozygous and segregates in subsequent generations. We measured growth rate in F_2 progeny generated from selfed F_1 , except for grLC4, where we used F_3 progeny. Owing to the trait's low heritability, QTL fine-mapping required fairly large sample sizes; including heterozygotes, we measured above-ground biomass from nearly 7,000 plants.

Plant growth conditions and fresh weight determination

We grew plants at a density of 337 plants per m^2 in 96-cell 'flats'. Each flat contained F_2 (or F_3) progeny from one of the advanced crosses. We planted seeds into damp, autoclaved Klasmann tray substrate (Klasmann-Deilmann GmbH) mixed with vermiculite at a 3:1 ratio. To the bottom of each flat, we added 22 g Osmocote Vegetable and Bedding 14-14-14 Plant Food (Scotts Company) as a slow-release fertilizer. We covered flats with clear plastic grow domes. Seeds stratified for 3 days at 6 °C in the dark. Afterwards, we moved the flats to environment-controlled climate chambers (York), adjusted to 10.5 h light at 22 °C and 60% relative humidity, and 12.5 h darkness at 16 °C and 80% relative humidity. Light and dark phases were both followed by a 0.5-h transition, simulating dusk and dawn, respectively. Illumination was provided by 4 out of 12 NH360/FLX/coated SUNLUX ACE bulbs (RQL) at a distance of 1.5 m from the shelves. Every day, four different light bulbs per shelf were used. Seeds germinated in 2–3 days, and the germination date of each seed was noted. For each flat, we included only those plants in the analyses that had germinated at the same day. We removed grow domes after 5 days under lights. In general, we grew two or three flats per advanced cross simultaneously and adjacent to one another. We rotated flats daily at noon, and shifted them from end to end to compensate for slight variation in temperature and light intensity in the grow room. We watered plants from below twice during growth and at the day before harvest with 1.5–2 l distilled water per flat to maintain turgor. We measured fresh weight of aerial plant parts 16 days after germination; we cut off shoots immediately above the roots, removed remaining soil with a paper towel, and weighed them to the nearest centigram. Plants weighing less than 0.01 g were excluded from subsequent analyses. Subsequently, we freeze-dried one or two leaves per plant, or, in case of small plants, the entire sample, for DNA extraction and genotyping.

Statistical analyses

We used Systat Version 10 (SPSS Inc.) for mixed model analysis of variance (ANOVA) of above ground biomass accumulation. We treated genotype as a fixed effect, and flat and flat $\times$ genotype interaction as random effects to obtain mean squares (MS) for genotype and flat $\times$ genotype in the model: fresh weight = constant + genotype + flat + flat $\times$ genotype. We calculated F ratios as $MS_{\text{genotype}}/MS_{\text{flat} \times \text{genotype}}$. In general, we performed two-tailed tests for growth rate differences except for the control cross grCL(6–9), which was designed to confirm presence and direction of the QTL identified in grCL7, and where we confirmed statistical significance by a one-tailed test. We used DnaSP 4.00 (ref. 22) for statistical tests of molecular population genetics.

See Supplementary Methods for details of genotyping and the At5g23170 sequence survey from 31 *Arabidopsis* accessions.

Received 22 November 2004; accepted 17 February 2005; doi:10.1038/nature03480.

- Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. & Hirschhorn, J. N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genet.* **33**, 177–182 (2003).
- Mackay, T. F. C. The genetic architecture of quantitative traits: lessons from *Drosophila*. *Curr. Opin. Genet. Dev.* **14**, 253–257 (2004).
- Steinmetz, L. M. *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330 (2002).
- Koornneef, M., Alonso-Blanco, C. & Vreugdenhil, D. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.* **55**, 141–172 (2004).
- Dilda, C. L. & Mackay, T. F. C. The genetic architecture of *Drosophila* sensory bristle number. *Genetics* **162**, 1655–1674 (2002).
- Kroymann, J. *et al.* A gene controlling variation in *Arabidopsis thaliana* glucosinolate composition is part of the methionine chain elongation pathway. *Plant Physiol.* **127**, 1077–1088 (2001).
- Kroymann, J., Donnerhacke, S., Schnabelrauch, D. & Mitchell-Olds, T. Evolutionary dynamics of an *Arabidopsis* insect resistance QTL. *Proc. Natl Acad. Sci. USA* **100**, 14587–14592 (2003).
- Lister, C. & Dean, C. Recombinant inbred lines for mapping RFLP and phenotypic markers. *Plant J.* **4**, 745–750 (1993).
- Becraft, P. W., Stinard, P. S. & McCarty, D. R. CRINKLY4: a receptor kinase with TNFR similarity, involved in maize epidermal differentiation. *Science* **273**, 1406–1409 (1996).
- Shiu, S. H. & Blecker, A. B. Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *Proc. Natl Acad. Sci. USA* **98**, 10763–10768 (2001).
- Becraft, P. W. Receptor kinase signaling in plant development. *Annu. Rev. Cell Dev. Biol.* **18**, 163–192 (2002).
- Zhou, J., Loh, Y. T., Bressan, R. A. & Martin, G. B. The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. *Cell* **83**, 925–935 (1995).

- Wright, S. I., Lauga, B. & Charlesworth, D. Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. *Mol. Biol. Evol.* **19**, 1407–1420 (2002).
- Tian, D., Araki, H., Stahl, E., Bergelson, J. & Kreitman, M. Signature of balancing selection in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **99**, 11525–11530 (2002).
- Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
- Schmid, K. J., Ramos-Onsins, S., Ringys-Beckstein, H., Weisshaar, B. & Mitchell-Olds, T. A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. *Genetics* **169**, 1601–1615 (2005).
- Haubold, B., Kroymann, J., Ratzka, A., Mitchell-Olds, T. & Wiehe, T. Recombination and gene conversion in a 170-kb genomic region of *Arabidopsis thaliana*. *Genetics* **161**, 1269–1278 (2002).
- Nordborg, M. *et al.* The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genet.* **30**, 190–193 (2002).
- Lauter, N. & Doebley, J. Genetic variation for phenotypically invariant traits detected in Teosinte: implications for the evolution of novel forms. *Genetics* **160**, 333–342 (2002).
- Gibson, G. & Dworkin, I. Uncovering cryptic genetic variation. *Nature Rev. Genet.* **5**, 681–690 (2004).
- Torii, K. U. *et al.* The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* **8**, 735–746 (1996).
- Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank S. Donnerhacke, D. Schnabelrauch, K. Eberhardt, K. Jünemann and A. Hirsch for help with plant care and technical assistance. This work was supported by the Max Planck Society, the DFG and the European Union.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.K. (kroymann@ice.mpg.de). Sequence data are deposited at EMBL under accession numbers AJ864968–AJ864998.

Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity

Li-Ru You¹, Fu-Jung Lin^{1*}, Christopher T. Lee^{1*}, Francesco J. DeMayo^{1,2}, Ming-Jer Tsai^{1,2} & Sophia Y. Tsai^{1,2}

¹Department of Molecular and Cellular Biology, ²Developmental Biological Program, Baylor College of Medicine, Houston, Texas 77030, USA

*These authors contributed equally to this work

Arteries and veins are anatomically, functionally and molecularly distinct. The current model of arterial–venous identity proposes that binding of vascular endothelial growth factor to its heterodimeric receptor—Flk1 and neuropilin 1 (NP-1; also called Nr1p)—activates the Notch signalling pathway in the endothelium, causing induction of ephrin B2 expression and suppression of ephrin receptor B4 expression to establish arterial identity^{1–4}. Little is known about vein identity except that it involves ephrin receptor B4 expression, because Notch signalling is not activated in veins; an unresolved question is how vein identity is regulated. Here, we show that COUP-TFII (also known as Nr2f2), a member of the orphan nuclear receptor superfamily, is specifically expressed in venous but not arterial endothelium. Ablation of COUP-TFII in endothelial cells enables veins to acquire arterial characteristics, including the expression of arterial markers NP-1 and Notch signalling molecules, and the generation of haematopoietic cell clusters. Furthermore, ectopic expression of COUP-TFII in endothelial cells results in the fusion of veins and arteries in transgenic mouse embryos. Thus, COUP-TFII has a critical role in repressing Notch signalling to maintain vein identity, which suggests that vein identity is under genetic control and is not derived by a default pathway.

Targeted disruption of the *COUP-TFII* gene showed that

COUP-TFII has an important role in angiogenesis and heart development⁵. To examine further the role of COUP-TFII in vascular development, we studied COUP-TFII expression using the *COUP-TFII/lacZ* 'knock-in' mouse model⁶. 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal) staining shows that COUP-TFII expression occurs in the venous endothelium at embryonic day (E)8.5 (data not shown). As development proceeds, high levels of COUP-TFII expression are detected in the endothelium of veins, including anterior cardinal vein, umbilical vein and vitelline vein, but not in arterial endothelial cells, including dorsal aorta, internal carotid artery and umbilical artery (Supplementary Fig. 1). To ensure that COUP-TFII is only expressed in venous endothelium, we used Tie2 (refs 7, 8) as an endothelial

marker. The co-localization of Tie2 with COUP-TFII in venous endothelium further confirmed that COUP-TFII is specifically expressed in the endothelium of cardinal (Fig. 1a) and umbilical veins (Fig. 1c). In contrast, all arterial endothelial cells in the dorsal aorta (Fig. 1a, b) and umbilical artery (Fig. 1c, d) show Tie2 expression, but not COUP-TFII expression. In addition, COUP-TFII is expressed in the smooth muscle cells surrounding arteries, as shown by co-localization of COUP-TFII with smooth muscle actin- α (SMA- α) (Fig. 1b, d).

To evaluate a possible role of *COUP-TFII* in the establishment of arterial-venous identity, we used a chimaera analysis. Both wild-type and mutant *COUP-TFII* embryonic stem (ES) cells containing the *ROSA26* allele were microinjected into wild-type mouse blastocysts to generate chimaeric embryos (Supplementary Fig. 2). If *COUP-TFII* has a cell-autonomous function in the formation of venous endothelium, mutant cells will not be able to contribute to endothelial formation. Thus, X-gal-positive cells will be excluded from venous endothelium. In contrast, if *COUP-TFII* has a non-cell-autonomous function in the formation of the endothelium of veins, mutant cells will be able to contribute, and X-gal-positive cells will be intermixed with wild-type cells. In control mice, X-gal-positive wild-type, ES-derived cells are seen in all cell lineages of the umbilical artery (Fig. 1e) and umbilical vein (Fig. 1f). In contrast, *COUP-TFII* mutant cells are excluded from venous endothelium (Fig. 1h) but not from arterial endothelium (Fig. 1g). In addition, *COUP-TFII* mutant cells are also excluded from the smooth muscle layer of arteries (Fig. 1g). These data indicate that *COUP-TFII* has a cell-autonomous function in the formation of venous endothelium and smooth muscle cells of arteries.

To test the involvement of *COUP-TFII* in the arterial-venous cell fate decision, we generated endothelial-specific knockout mice by crossing floxed (*loxP* sites were inserted in the 3' and 5' untranslated region of the *COUP-TFII* gene locus) *COUP-TFII* mice⁶ with the endothelial-specific *Tie2-Cre* mouse line⁹ that has been used to efficiently ablate ephrin B2 (*Efnb2*)² and NP-1 (ref. 4) in endothelial cells. Because COUP-TFII is not expressed in the endothelium of arteries, COUP-TFII expression is only ablated in the endothelium of veins in these mutants (Supplementary Fig. 3). No *Tie2-Cre/+; COUP-TFII^{lox/lox}* mice were recovered at birth, suggesting that inactivation of COUP-TFII in venous endothelial cells results in embryonic lethality by E12. The endothelial-specific *COUP-TFII* conditional null mutants exhibit a variety of vascular defects, including widespread haemorrhage and thin and well-dilated vessels (Supplementary Fig. 4); however, the collapsed or missing anterior cardinal veins displayed by the *COUP-TFII* null embryos⁵ are not observed in the endothelial-specific knockout embryos.

To determine whether loss of COUP-TFII results in the acquisition of arterial characteristics in mutant veins, we examined the expression of Notch1 (refs 1, 10, 11) and its ligand jagged 1 (*Jag1*)^{12,13}. Expression of *Jag1* was detected in endothelial and smooth muscle cells of arteries of both control and mutant (*Tie2-Cre/+; COUP-TFII^{lox/lox}*) mice (Fig. 2a, b). In contrast, *Jag1*, which was not expressed in control veins, was ectopically expressed in the veins of mutant mice (Fig. 2b). Similarly, Notch1, which was not expressed in control veins (Fig. 2c), was ectopically expressed in the endothelium of mutant veins (Fig. 2d). We next examined the expression of *Efnb2* (refs 14, 15), a downstream arterial endothelial marker that is induced by Notch signalling. Using RNA *in situ* hybridization, we found that *Efnb2* was expressed in the arteries (Fig. 2e, f) of control and mutant mice, but not in the veins of control mice (Fig. 2e). However, ectopic expression of *Efnb2* was observed in the veins of mutant mice (Fig. 2f). To assess whether COUP-TFII regulates factors upstream of the Notch signalling pathway^{3,16,17}, we analysed the expression of NP-1. NP-1 is highly expressed in arteries but is barely detectable in the veins of

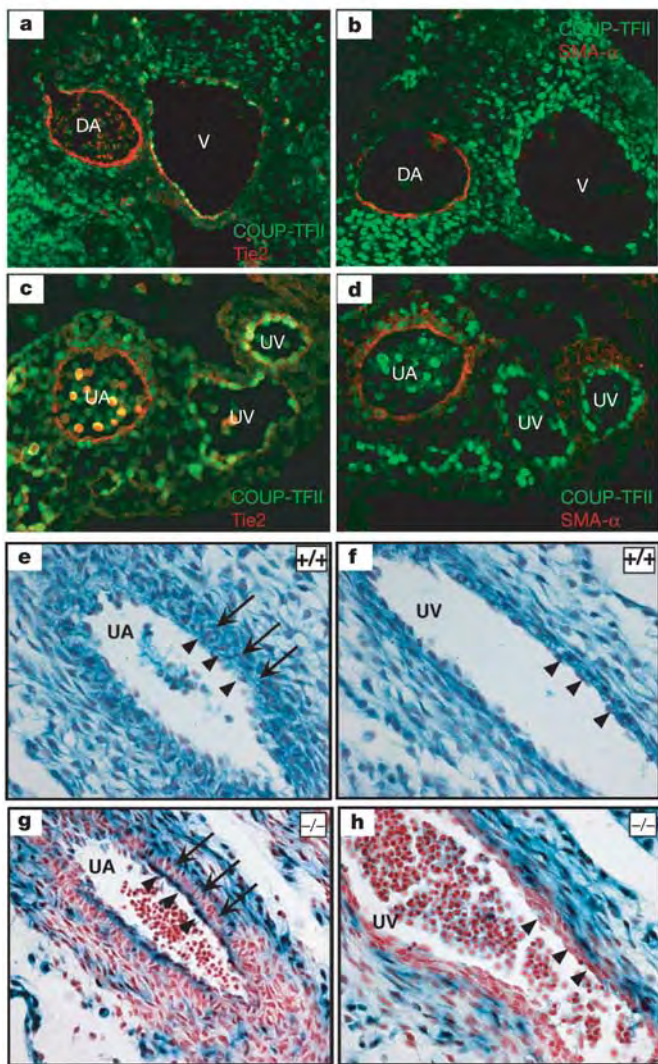


Figure 1 Differential expression pattern and chimaera analysis of COUP-TFII in the vasculature. **a–d**, COUP-TFII (green) and Tie2 (red) expression in the endothelial cells of the cardinal vein at E11.5 (V; **a**), umbilical vein (UV) at E9.5 (**c**), but not in the dorsal aorta (DA; **a**) or umbilical artery (UA; **c**). COUP-TFII and SMA- α (red) expression co-localized in the smooth muscle cells surrounding the dorsal aorta (**b**) and umbilical artery (**d**). **e–h**, COUP-TFII is required cell autonomously in arterial and venous development. The X-gal-positive cells are seen in all cell lineages in the umbilical artery and umbilical vein (**e**, **f**) using chimaera analysis (Supplementary Fig. 1). In contrast, in E12.5 chimaeras generated from *lacZ*-positive *COUP-TFII*^{+/+} ES cells, *COUP-TFII* mutant cells (blue) are able to contribute to endothelial cells of the artery (**g**, arrowheads), but are excluded from the endothelial cell layer of umbilical vein (**h**, arrowheads) and the subendothelial smooth muscle layer around it (**g**, arrows).

control mice (Fig. 2g). In veins of mutant mice, NP-1 expression is greatly enhanced relative to control veins (Fig. 2g, h), suggesting that COUP-TFII acts at or upstream of NP-1 to suppress Notch signalling in the endothelium of veins. Therefore, COUP-TFII is the first transcription factor to be implicated in regulating vein identity.

The lack of arterial–venous fusion in *COUP-TFII*-deficient embryos supports the theory that mutant veins might not have fully lost their venous identity. Because the *notch–gridlock* (*grl*) signalling pathway guides arterial fate in zebrafish¹⁸, we asked whether mouse homologues of *grl* (*Hey1*¹⁹ genes) might be induced in mutant veins. Using *in situ* hybridization, we found that *Hey1* is expressed in the arteries of both control and mutant embryos (Fig. 2i, j). However, *Hey1* is ectopically expressed only in mutant veins, but its level of expression is not as high as compared with arterial expression (Fig. 2i, j). Under normal conditions, *Hey1*/*grl* signalling suppresses the expression of the vein marker *Ephb4* (ref. 18). We next examined whether expression of *Ephb4* is suppressed in the veins of mutant mice. In the veins of *COUP-TFII* mutant mice, levels of endothelial *Ephb4* are only slightly reduced (Fig. 2k, l). It is possible that the low level of *Hey1* expression in mutant veins is not sufficient to suppress effectively *Ephb4*

expression, and thus it fails to fully convert veins to an arterial fate. This could explain why no arterial–venous fusion is observed in the *COUP-TFII*-deficient mice.

A unique feature of functional arterialization of a vessel is the formation of haematopoietic cell clusters at the AGM (aorta–gonad–mesonephros) region, which have been characterized in chick, mouse and human embryos^{20,21}. These clusters develop at the ventral side of the dorsal aorta and in the vitelline and umbilical arteries in the AGM region^{22,23}; they have never been observed in association with veins. To assess whether mutant veins behaved in a similar way to arteries of the AGM region, we determined whether haematopoietic cell clusters appeared in the veins of *COUP-TFII* mutant mice. Budding haematopoietic cell clusters were observed within mutant veins in several regions, including the head vein (Fig. 3a, d), the cardinal vein (Fig. 3b, e) and the umbilical vein (Fig. 3c, f), suggesting that the mutant veins can acquire functional features of arteries in various locations not restricted to the AGM region.

To determine whether the ectopic venous haematopoietic cell clusters possessed the same characteristics as haematopoietic cell clusters normally detected in arteries of the AGM region, markers of

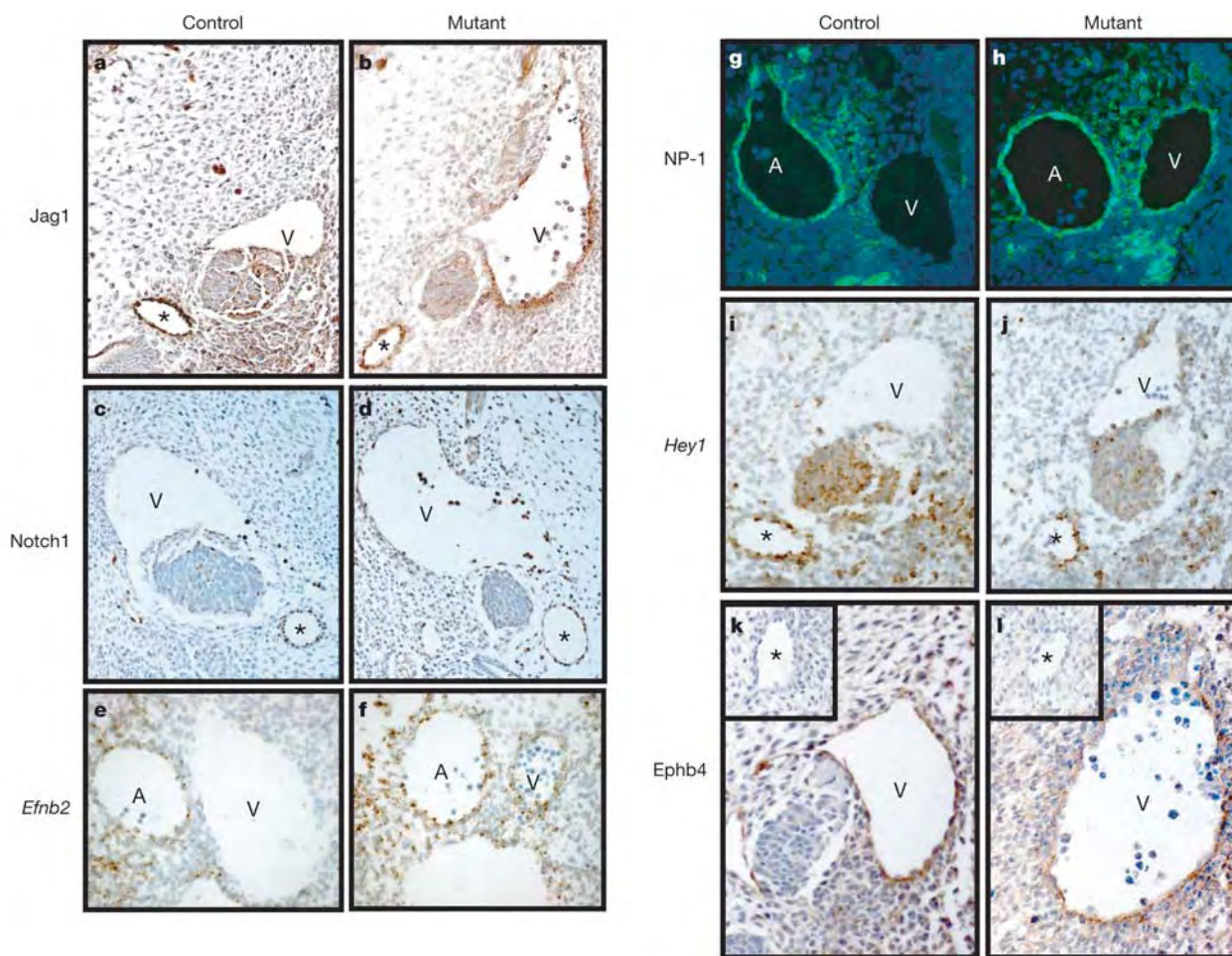


Figure 2 Arterial markers are ectopically expressed, whereas venous marker expression is reduced in veins of *COUP-TFII* mutant mice. **a–d**, Jag1 (**a, b**) and Notch1 (**c, d**) are ectopically expressed in the veins (V) of *COUP-TFII* mutant mice (*Tie2-Cre*⁺; *COUP-TFII*^{flax/flax}) (**b, d**) but not in control (*Tie2-Cre*⁺; *COUP-TFII*^{flax/+}) veins (**a, c**); they are also detected in transverse sections of E11.5 arteries. **e, f**, *In situ* hybridization shows ectopic *Efnb2* expression in mutant (**f**) but not control (**e**) vein, whereas it is expressed in the arteries (A) of both E11.5 control and mutant mice at the common cardinal vein level. **g, h**, NP-1 (green) is highly expressed in artery but is barely detectable in the vein of a

control E10.5 mouse (**g**). NP-1 expression is greatly enhanced in the vein of a E10.5 mutant mouse (**h**). Nuclei were counterstained with DAPI (blue). **i, j**, *In situ* hybridization shows ectopic expression of *Hey1* in mutant (**j**) but not control (**i**) vein, whereas *Hey1* is expressed in both control and mutant arteries at the E11.5 head vein level (**i, j**). **k, l**, *Ephb4* immunostaining of E11.5 transverse sections at the head vein level shows that *Ephb4* expression is slightly reduced in mutant vein (**l**) as compared with control vein (**k**), whereas no expression of *Ephb4* is detected in the internal carotid artery of control or mutant mice (inset; marked by asterisk in **a–d, i–l**).

haematopoietic cell clusters, including CD34, SMA- α and von Willebrand factor (vWF), were measured. CD34 and vWF are normally expressed in the endothelium of both arteries and veins in control and mutant mice (asterisks in Fig. 3g, i, j, l). However, CD34- and vWF-positive cell clusters were detected in veins of mutant mice (Fig. 3j, l) but not in controls (Fig. 3g, i). Similarly, SMA- α is expressed in smooth muscle cells surrounding the arteries of both control and mutant mice (asterisks in Fig. 3h, k), but SMA- α positive haemogenic sites are only detected in mutant veins (Fig. 3k) and not controls (Fig. 3h). The identities of these haematopoietic cell clusters were further confirmed using immunostaining for anti-CD45 and anti-c-kit (CD117) antibodies (Supplementary Fig. 5). These results demonstrate that veins of *COUP-TFII* mutant mice acquire arterial function and serve as sites for intra-embryonic haematopoiesis. The haematopoietic stem cell niche regulates stem cell fate and cell numbers^{24,25}. Activation of the Notch1–Jag1

signalling pathway increases haematopoietic stem cell numbers in the haematopoietic stem cell niche of adult tissue^{24,25}. The activation of Notch1 and ectopic expression of Jag1 in the veins of *COUP-TFII* mutant mice suggests that disruption of *COUP-TFII* in the venous endothelium may affect the regulatory microenvironment in the haematopoietic stem cell niche.

To support further the notion that *COUP-TFII* suppresses both NP-1 expression and the Notch signalling pathway, endothelial-specific overexpressed *COUP-TFII* transgenic embryos were generated (*Tie2-COUP-TFII*; Fig. 4a). The yolk sac of transgenic embryos exhibits defects in angiogenesis (Fig. 4c, d), whereas transgenic embryos show pericardium effusion, fail to turn (Fig. 4e, f) and start to die from E10–E11.5. Although the transgenic embryo represents an independent insertion of the transgene, these phenotypes were highly reproducible. The transgenic embryos exhibit large, fused, disorganized vessels without arterial–venous distinction (Fig. 4h).

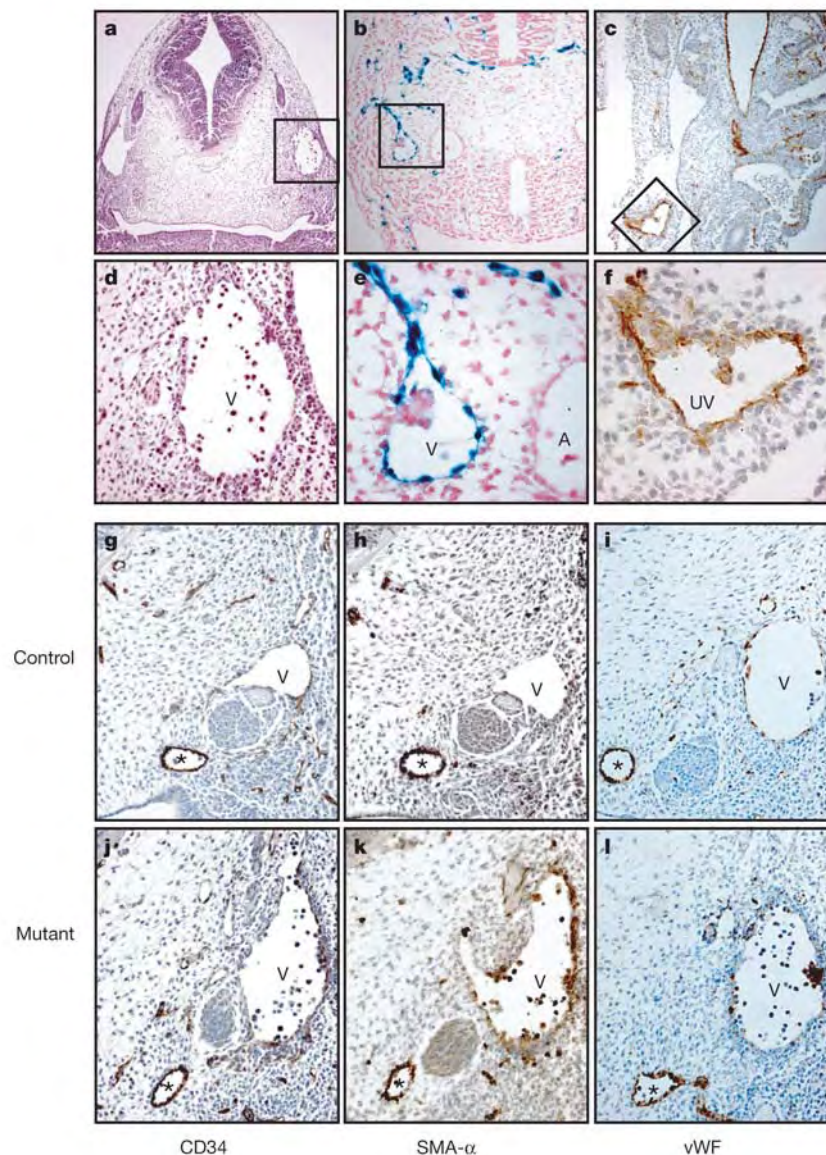


Figure 3 Ectopic formation of haematopoietic cell clusters in *COUP-TFII* mutant venous endothelium. The haematopoietic cell clusters are detected in veins of mutant mice (*Tie2-Cre/+; COUP-TFII^{fllox/fllox}*); they are not restricted to the AGM region. **a, d**, A haematoxylin- and eosin-stained E11.5 transverse section shows a haematopoietic cell cluster budding from the head vein in a *COUP-TFII* mutant mouse (**a**). Panel **d** is a higher-magnification image of the boxed region in **a**. **b, e**, An X-gal-stained E10.5 transverse

section shows the cluster budding from the cardinal vein (V; **b**). A higher-magnification image is shown in **e**. **a**, aorta. **c, f**, A CD34-positive cluster detected in E10.5 umbilical vein (UV; **c**). A higher-magnification image is shown in **f**. **g–l**, Haematopoietic cell cluster-specific markers CD34 (**g, j**), SMA- α (**h, k**) and vWF (**i, l**) are detected in the veins of mutant mice (*Tie2-Cre/+; COUP-TFII^{fllox/+}*) but not in the veins of control mice. The internal carotid artery is marked by an asterisk.

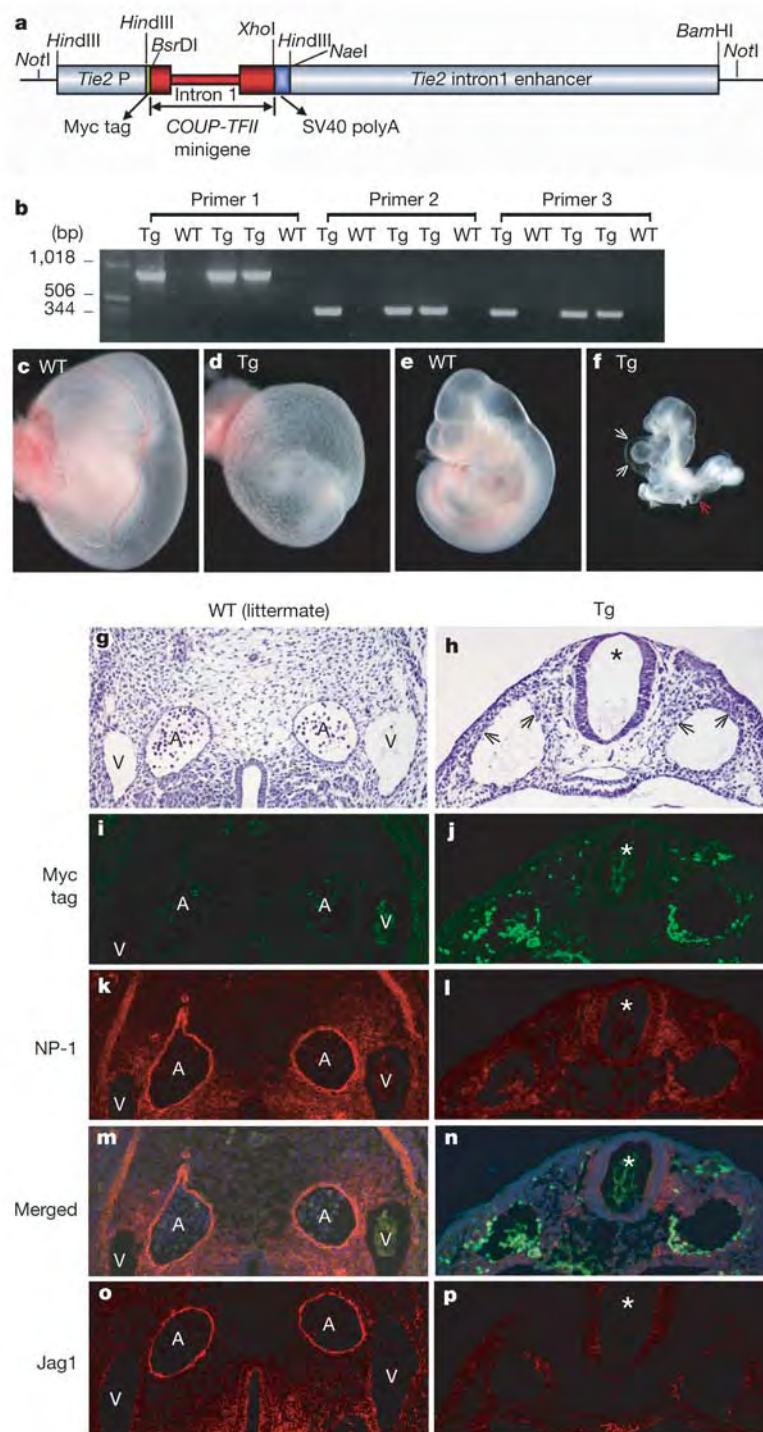


Figure 4 Mis-expression of COUP-TFII in the endothelium. **a**, Diagram of the *Tie2-COUP-TFII* transgene (Tg) construct. **b**, The transgenic embryos were confirmed by PCR genotyping. **c–f**, Defective angiogenesis in the transgenic embryos. No big blood vessels can be detected in the yolk sac of transgenic embryos (**d**), whereas large vessels are noted in controls (**c**). The transgenic embryo is growth retarded, with pericardium effusion (white arrows), and fails to turn (**f**, red arrow) in comparison to a control embryo (**e**). **g, h**, Haematoxylin- and eosin-stained transverse sections show that the transgenic embryo (**h**)

is smaller than the control (**g**). Also, veins and arteries in the transgenic embryo are fused into a single entity (arrows in **h**). **i, j**, Expression of the transgene is detected in the large, fused endothelium by Myc-tagged antibody (green, **j**). Myc signals were not detected in a wild-type control (**i**). **k–p**, NP-1 (**k–n**) and Jag1 (**o, p**) are expressed at high levels in control arteries (red, **k, o**) but are barely detectable in the endothelium of transgenic mice (red, **l, p**). **m, n**, Merged images of *COUP-TFII* transgene and NP-1. Nuclei were counterstained by DAPI (blue). Asterisks in **h, j, l, n, p** denote neural tube.

These phenotypes resemble those displayed by NP-1 (ref. 17) and Notch1 (ref. 10) null mutants, suggesting that NP-1 might be downregulated in the transgenic embryos. Expression of NP-1 is greatly reduced in the defective vessels of the COUP-TFII transgenic embryos (Fig. 4l, n), but is strongly expressed in the arteries of

wild-type littermates (Fig. 4k, m). The reduced expression of NP-1 is coincident with the ectopic expression of the Myc-tagged COUP-TFII transgene, as visualized by Myc-tagged antibody immunostaining (Fig. 4j, n). Expression of Jag1, an arterial marker (Fig. 4o), is significantly reduced in the transgenic embryos (Fig. 4p). The

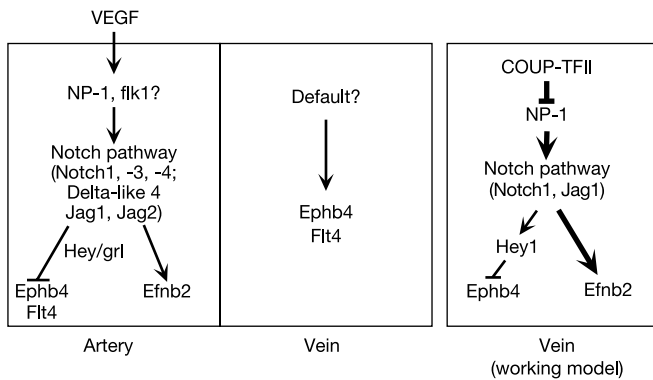


Figure 5 COUP-TFII maintains vein identity. Existing model of arterial–venous differentiation²⁷ (adapted and modified from ref. 27) and our working model of the regulation of vein fate decision by COUP-TFII. Under normal conditions, COUP-TFII suppresses NP-1 to downregulate Notch1, Jag1 and Efnb2 in the endothelium of veins. In the absence of COUP-TFII, aberrant expression of arterial markers coupled with a slight reduction in the vein marker Ephb4 in mutant veins results in the partial conversion of mutant vein to artery (that is, veins acquire the molecular and functional characteristics of arteries). Our model suggests that vein fate decision is highly complex and probably involves the participation of COUP-TFII with as-yet-unidentified factors to regulate this complex process. VEGF, vascular endothelial growth factor.

vein-like nature of the defective vessels is further confirmed by their positive staining with Ephb4 antibody (data not shown). Taken together, these results indicate that COUP-TFII functions upstream of NP-1 and Notch signalling to regulate arterial–venous differentiation.

On the basis of our findings, we propose a working model for the role of COUP-TFII in establishing arterial–venous identity (Fig. 5). During normal development, COUP-TFII in venous endothelium suppresses NP-1 and inhibits Notch signalling and the expression of arterial-specific genes. Without the effects of NP-1 and Notch signalling, Ephb4, Flt4 and other factors important for the differentiation of the vein are expressed and vein identity is maintained. Removal of COUP-TFII from the venous endothelium activates the expression of arterial markers NP-1, Jag1, Notch1 and Efnb2, leading to the acquisition of arterial functions and the induction of haematopoietic cell clusters. Notably, the acquisition of these arterial characteristics is not sufficient to convert mutant veins fully into arteries, because the expression of Ephb4 is not totally lost, and arterial–venous fusion is not detected in the mutant mice. In contrast, mis-expression of COUP-TFII in the endothelium of arteries suppresses the expression of NP-1 and the downstream arterial markers, changing the artery into a more vein-like vessel. Thus, the establishment and maintenance of arterial–venous identity is tightly regulated. These findings have broadened our understanding of the developmental regulation of arterial–venous identity, and elucidate a new role for the COUP-TFII transcription factor in venous cell fate decision. □

Methods

Conditional knockout mice and genotyping

Generation of *COUP-TFII*/*lacZ* knock-in mice and floxed *COUP-TFII* mice has been described⁶ (see also Supplementary Fig. 3). Using a homologous recombination strategy, we inserted three *loxP* sites into the *COUP-TFII* locus. Upon Cre-mediated recombination between the *loxP* sites flanking the genomic DNA of *COUP-TFII*, *lacZ* was turned on under the control of the *COUP-TFII* promoter. The staining pattern of X-gal enabled us to follow the recombination event. Endothelial-specific *Tie2-Cre* transgenic mice were provided by M. Yanagisawa. To generate the endothelial-specific knockout mice, the floxed *COUP-TFII* mice were first crossed with the *Tie2-Cre* line to generate *Tie2-Cre*⁺; *COUP-TFII*^{lox/lox} compound heterozygote mice. The male compound heterozygote was then crossed to *COUP-TFII*^{lox/lox} mice to obtain *Tie2-Cre*⁺; *COUP-TFII*^{lox/lox} mice, to ablate *COUP-TFII* in the endothelium.

Generation of transgenic embryos and genotyping

The 2.1-kilobase (kb) promoter (*HindIII*–*HindIII*) and 10.5-kb first intron sequences (*NaeI*–*BamHI*) of the murine *Tie2* gene from bacterial artificial chromosome (BAC) RP23-416H4 (Invitrogen) were subcloned into pBSISK(-) (Stratagene). The resulting plasmids were called pBSTie2-2 and pBSTie2-10.5, respectively. Simian virus 40 polyA signal sequence from pCMV-Tag 3C (Stratagene) was inserted into the *HindIII* site in pBSTie2-10.5 to generate pBSpolyATie2-10.5. To differentiate the transgene from endogenous *COUP-TFII*, Myc tag was synthesized and ligated in frame to the *COUP-TFII* minigene (comprising *COUP-TFII* complementary DNA with intron 1, 2.9 kb) and placed downstream of the 2.1-kb *Tie2* promoter. The *XhoI* fragment containing the *Tie2* 2.1-kb promoter and the Myc-tagged *COUP-TFII* minigene was removed from the resultant plasmid and cloned into pBSpolyATie2-10.5. The final Tg(*Tie2-COUP-TFII*) transgene, comprising the *Tie2* promoter, Myc-tagged *COUP-TFII* minigene, simian virus 40 polyA and *Tie2* intron 1 enhancer, was excised from the vector and used to generate transgenic mice. Genomic DNA from the yolk sac of transgenic embryos and wild-type littermates was confirmed by polymerase chain reaction (PCR). Only DNA from the transgenic embryo contained positive PCR products. Primer 1: *COUP-TFII*/*HindIII*-f (forward, 5'-CTGTCTGATGTAGCCCCATGTGG-3') and *Tie2*-533r (reverse, 5'-CCATGTGTAAAGTGGAGAGTCTC-3'); primer 2: *Tie2*P-2628-f (forward, 5'-GTAACAAGATCCCTCTGCC-3') and *COUP*-170-r (reverse, 5'-CACGTGCTGACTACCATTGC-3'); primer 3: *Tie2*P-2628-f and Myc-BsrD1-r (reverse, 5'-CATTGCCAACAGATCCTCTTC-3').

X-gal staining

For frozen sections, embryos were dissected and fixed in 2% paraformaldehyde (PFA)/PBS, pH 7.4 for 30 min at 4°C, washed with PBS, cryo-protected by 20% sucrose/PBS and embedded in OCT compound. The sections were stained with X-gal, as described previously²⁶, and counterstained with Nuclear Fast red.

Histology, immunohistochemistry and *in situ* hybridization

Embryos for histological analysis were dissected and fixed in 4% PFA/PBS overnight. Fixed embryos were processed, embedded in paraffin, sectioned and stained with haematoxylin and eosin as previously described⁵. For the immunohistochemistry analysis, antibodies to COUP-TFII (1:6,000), CD34, CD45 and c-kit (CD117) (BD Pharmingen), smooth muscle α -actin (SMA- α , Sigma, clone 1A4, 1:2,000), vWF (DAKO, 1:1,500), Notch1 (Cell Signalling, 1:150), Ephb4 (R&D systems, 1:200), Jag1 (Santa Cruz, 1:1,000), rNeuropilin-1 (R&D systems, 1:600) and Myc tag (9E10) (Santa Cruz, 1:500) were used. Biotinylated secondary antibodies (1:300, Vector) and the avidin/streptavidin-based detection system (Vectastain elite ABC kit, Vector) were used, developed in 3,3'-diaminobenzidine tetrahydrochloride (DAB) (Vector), and counterstained with haematoxylin. For double-immunostaining, Alexa fluor-488 and -594-conjugated tyramide signal amplification kits (Molecular Probes) were applied to amplify the signal. For non-radioactive *in situ* hybridization, antisense messenger RNA probes for *Efnb2* and *Hey1* (nucleotides 794–2181 of NM_010423) were prepared according to protocols provided by the manufacturer (Roche). Digoxigenin-labelled anti-sense RNA and sense control RNA (100–200 ng ml⁻¹) were incubated with the slide at 65°C in hybridization buffer (Ambion) overnight. Horseradish peroxidase conjugated anti-digoxigenin (DAKO) antibody was used to detect digoxigenin-labelled RNA. Biotinyl-tyramide (DAKO) was applied to amplify the signal. The avidin/streptavidin-based detection system (Vectastain elite ABC kit, Vector) was used, developed in DAB (Vector) and counterstained with haematoxylin.

Received 27 December 2004; accepted 28 February 2005; doi:10.1038/nature03511.

- Lawson, N. D. *et al.* Notch signaling is required for arterial–venous differentiation during embryonic vascular development. *Development* **128**, 3675–3683 (2001).
- Gerety, S. S. & Anderson, D. J. Cardiovascular ephrinB2 function is essential for embryonic angiogenesis. *Development* **129**, 1397–1410 (2002).
- Lawson, N. D., Vogel, A. M. & Weinstein, B. M. *Sonic hedgehog* and *vascular endothelium growth factor* act upstream of Notch pathway during arterial endothelial differentiation. *Dev. Cell* **3**, 127–136 (2002).
- Gu, C. *et al.* Neuropilin-1 conveys Semaphorin and VEGF signaling during neural and cardiovascular development. *Dev. Cell* **5**, 45–57 (2003).
- Pereira, F. A., Qiu, Y. H., Chou, G., Tsai, M.-J. & Tsai, S. Y. The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* **13**, 1037–1049 (1999).
- Takamoto, N. *et al.* COUP-TFII is essential for radial and anterior–posterior patterning of the stomach. *Development* **132**, 2179–2182 (2005).
- Sato, T. N. *et al.* Distinct roles of the receptor kinase Tie-1 and Tie-2 in blood vessel formation. *Nature* **376**, 70–74 (1995).
- Suri, C. *et al.* Requisite role of angiopoietin-1, a ligand for TIE2 receptor, during embryonic angiogenesis. *Cell* **87**, 1171–1180 (1996).
- Kisanuki, Y. Y. *et al.* Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis *in vivo*. *Dev. Biol.* **230**, 230–242 (2001).
- Huppert, S. S. *et al.* Embryonic lethality in mice homozygous for a processing-deficient allele of Notch1. *Nature* **405**, 966–970 (2000).
- Krebs, L. T. *et al.* Notch signaling is essential for vascular morphogenesis in mice. *Genes Dev.* **14**, 1343–1352 (2000).
- Xue, Y. *et al.* Embryonic lethality and vascular defects in mice lacking the Notch ligand Jagged1. *Hum. Mol. Genet.* **8**, 723–730 (1999).
- Villa, N. *et al.* Vascular expression of Notch pathway receptors and ligands is restricted to arterial vessels. *Mech. Dev.* **108**, 161–164 (2001).
- Wang, H. U., Chen, Z.-F. & Anderson, D. J. Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. *Cell* **93**, 741–753 (1998).
- Gerety, S. S., Wang, H. U., Chen, Z.-F. & Anderson, D. J. Symmetrical mutant phenotypes of the

- receptor *EphB4* and its specific transmembrane ligand *ephrin-B2* in cardiovascular development. *Mol. Cell* **4**, 403–414 (1999).
16. Shalaby, F. *et al.* Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature* **376**, 62–66 (1995).
 17. Kawasaki, T. *et al.* A requirement for neuropilin-1 in embryonic vessel formation. *Development* **126**, 4895–4902 (1999).
 18. Zhong, T. P., Childs, S., Leu, J. P. & Fishman, M. C. Gridlock signalling pathway fashions the first embryonic artery. *Nature* **414**, 216–220 (2001).
 19. Fischer, A., Schumacher, N., Maier, M., Sendtner, M. & Gessler, M. The Notch target genes *Hey1* and *Hey2* are required for embryonic vascular development. *Genes Dev.* **18**, 901–911 (2004).
 20. Garcia-Porrero, J. A., Godin, I. E. & Dieterlen-lievre, F. Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. *Anat. Embryol. (Berl.)* **192**, 427–435 (1995).
 21. Tavian, M. *et al.* Aorta-associated CD34⁺ hematopoietic cells in the early human embryo. *Blood* **87**, 67–72 (1996).
 22. Medvinsky, A. L., Samoylina, N. L., Muller, A. M. & Dzierzak, E. A. An early pre-liver intraembryonic source of CFU-S in the developing mouse. *Nature* **364**, 64–67 (1993).
 23. Godin, I. E., Garcia-Porrero, J. A., Coutinho, A., Dieterlen-lievre, F. & Macros, M. A. Para-aortic splanchnopleura from early mouse embryos contains B1a cell progenitors. *Nature* **364**, 67–69 (1993).
 24. Zhang, J. *et al.* Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425**, 836–841 (2003).
 25. Calvi, L. M. *et al.* Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425**, 841–846 (2003).
 26. Hogan, B. L. M., Beddington, R. S. P., Costantini, E. & Lacy, E. *Manipulating the Mouse Embryo, a Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1994).
 27. Adams, R. H. Molecular control of arterial-venous blood vessel identity. *J. Anat.* **202**, 105–112 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank W. Qian and C. Yang for technical assistance; M. Yanagisawa for providing the *Tie2-Cre* transgenic mice; Y. Furuta for providing the *Efnb2* construct for RNA *in situ* hybridization; and F. Petit for the *COUP-TFII* minigene construct. We also thank L.-Y. Yu-Lee and H. J. Bellen for discussions and critical reading of the manuscript. This work was supported by NIH grants to S.Y.T. and M.-J.T.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.Y.T. (stsai@bcm.tmc.edu) or M.-J.T. (mtsai@bcm.tmc.edu).

LPA₃-mediated lysophosphatidic acid signalling in embryo implantation and spacing

Xiaoqin Ye^{1*}, Kotaro Hama^{2*}, James J. A. Contos⁴, Brigitte Anliker¹, Asuka Inoue¹, Michael K. Skinner⁵, Hiroshi Suzuki^{3,6}, Tomokazu Amano³, Grace Kennedy¹, Hiroyuki Arai², Junken Aoki² & Jerold Chun¹

¹Department of Molecular Biology, Helen L. Dorris Child and Adolescent Neuropsychiatric Disorder Institute, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

²Graduate School of Pharmaceutical Sciences, and ³Department of Developmental Medical Technology (Sankyo), Graduate School of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

⁴Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington 98109-1024, USA

⁵Center for Reproductive Biology, School of Molecular Bioscience, Washington State University, Pullman, Washington 99164-4231, USA

⁶National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Obihiro 080-8555, Japan

* These authors contributed equally to this work

Every successful pregnancy requires proper embryo implantation. Low implantation rate is a major problem during infertility treatments using assisted reproductive technologies¹. Here we report a newly discovered molecular influence on implantation through the lysophosphatidic acid (LPA) receptor LPA₃ (refs 2–4). Targeted deletion of LPA₃ in mice resulted in significantly reduced litter size, which could be attributed to delayed

implantation and altered embryo spacing. These two events led to delayed embryonic development, hypertrophic placentas shared by multiple embryos and embryonic death. An enzyme demonstrated to influence implantation, cyclooxygenase 2 (COX2) (ref. 5), was downregulated in LPA₃-deficient uteri during pre-implantation. Downregulation of COX2 led to reduced levels of prostaglandins E₂ and I₂ (PGE₂ and PGI₂), which are critical for implantation¹. Exogenous administration of PGE₂ or carbaprostacyclin (a stable analogue of PGI₂) into LPA₃-deficient female mice rescued delayed implantation but did not rescue defects in embryo spacing. These data identify LPA₃ receptor-mediated signalling as having an influence on implantation, and further indicate linkage between LPA signalling and prostaglandin biosynthesis.

Multiple factors can adversely affect successful pregnancy. Two of these factors are failed synchronization between embryonic and endometrial development during implantation and occurrence of multiple gestations (especially monochorionic gestation), which can result in fetal demise^{1,6–9}. These factors are particularly important for the clinical success and efficacy of assisted reproductive technologies. One molecular factor that has been previously implicated in female reproduction is the small, bioactive phospholipid LPA¹⁰. LPA has a range of influences that are mediated by at least four G-protein-coupled receptors, LPA_{1–4} (ref. 2). Deletion of LPA₁ and LPA₂ in mice revealed roles for these receptors in neural development, craniofacial formation, neuropathic pain and altered cellular signalling, but without obvious effects on female reproduction^{11–14}. These results suggested that LPA signalling in female reproduction might be mediated by other LPA receptors including LPA₃ (formerly known as Edg7) (refs 3, 4), LPA₄ (ref. 15), unidentified LPA receptor(s), or possibly non-receptor pathways. Towards identifying LPA-dependent mechanisms affecting reproduction, we targeted LPA₃ for deletion. LPA₃ is a receptor with distinct signalling properties and a preference for unsaturated LPA species^{2–4}.

Functional deletion of LPA₃ was achieved by replacing a fragment covering the untranslated region and the start codon in exon 2 with a neomycin-resistance gene in reverse orientation in R1 embryonic stem cells (Supplementary Figs 1 and 2). The LPA₃-deficient mice were born with normal mendelian frequency without sexual bias (Supplementary Table 1), and appeared grossly normal (data not shown). However, LPA₃-deficient females produced litter sizes of less than 50% compared with that of wild-type and LPA₃ heterozygote controls (Supplementary Table 2), and showed a statistically significant prolongation of pregnancy (20.9 ± 0.5 days versus 19.4 ± 0.7 days in wild-type and LPA₃ heterozygote controls, *P* < 0.05). These phenotypes were independent of stud genotype, indicating defects in female reproduction.

Towards determining whether LPA₃ deletion might directly affect the female reproductive system, expression patterns of LPA₃ messenger RNA were assessed using polymerase chain reaction with reverse transcription (RT–PCR) and *in situ* hybridization. RT–PCR revealed the presence of LPA₃ mRNA in oviduct, placenta and uterus but not in ovary and eggs (unfertilized eggs and fertilized eggs from one cell to pre-implantation blastocyst; data not shown). Within the uterus, LPA₃ mRNA expression was upregulated during postnatal development and varied during the oestrous cycle (Supplementary Fig. 3a, b). Notably, LPA₃ mRNA levels increased during early pregnancy, peaking around embryonic day 3.5 (E3.5) then returning to basal levels from E4.5 through to the end of pregnancy (Fig. 1a). RT–PCR of microdissected E3.5 uterine tissue and *in situ* hybridization indicated that LPA₃ mRNA expression was confined to the luminal endometrial epithelium at E3.5 (Fig. 1b–d; see also Supplementary Fig. 3c). These data suggested that LPA₃ loss of function resulting in reduced litter sizes could involve direct effects on the female reproductive system.

To explore this possibility, we examined major events in female reproduction: from ovulation through to decidualization. No

significant differences were observed in superovulation, fertilization or decidualization between wild-type or LPA₃ heterozygote controls and LPA₃-deficient female littermates. No significant differences in blastocyst number or developmental stage, isolated from E3.5 uteri, were observed between control and LPA₃-deficient females. These data indicated no obvious defects in ovulation, ovum transportation and blastocyst development in LPA₃-deficient female mice (Fig. 1e; see also Supplementary Fig. 4).

In contrast, embryo implantation studies identified clear phenotypic changes in LPA₃-deficient dams: delayed implantation and

altered positioning/crowding of embryos. By E4.5, implantation sites identifiable by Evans blue labelling in control animals were absent in uteri of LPA₃-deficient dams. Implantation sites became detectable at E5.5 in uteri of LPA₃-deficient mice (Fig. 1f, g)¹⁶. The number of pre-implantation blastocysts recovered from E4.5 LPA₃-deficient uteri was comparable to that from E3.5 control and LPA₃-deficient uteri (Fig. 1e), indicating that delayed implantation in uteri of LPA₃-deficient mice was due to extra-embryonic influences of LPA signalling. In addition to delayed implantation, LPA₃-deficient uteri had a reduced number of implantation sites compared with that in the control uteri, despite the fact that comparable numbers of blastocysts were available for implantation. The implantation sites in the LPA₃-deficient uteri were crowded/clustered in the uterine segments proximal to the cervix (Fig. 1g). This aberrant crowding of embryos in uteri of LPA₃-deficient mice was further demonstrated by findings at later gestational stages. At E10.5, 44% of implantation sites in LPA₃-deficient uteri contained two to four embryos (averaging 1.65 live embryos per implantation site) (Fig. 2a, b). At E18.5, 28% of placentas were shared by two to

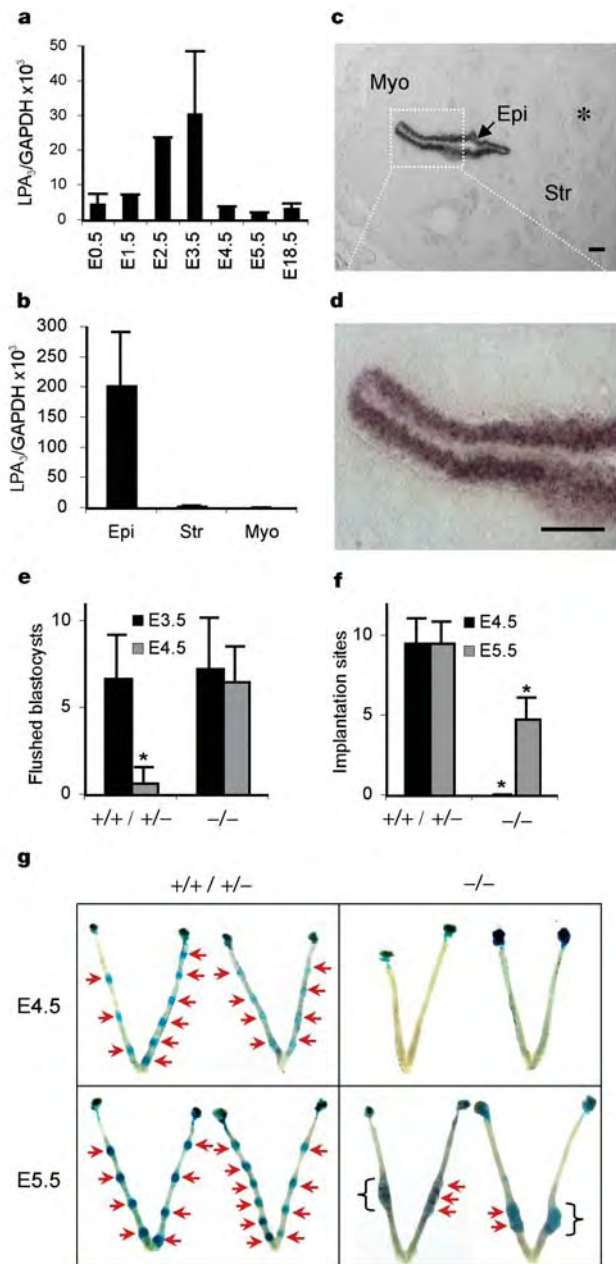


Figure 1 LPA₃ mRNA expression in wild-type uterus and effects of LPA₃ deficiency on implantation. **a, b**, Quantification of uterine LPA₃ mRNA during pregnancy (**a**) and in E3.5 luminal endometrial epithelium (Epi), stroma (Str) and myometrium (Myo) (**b**). **c, d**, *In situ* localization of LPA₃ in E3.5 wild-type uterus. The asterisk indicates glandular endometrial epithelium. Scale bars: 100 μ m. **e**, Number of flushed blastocysts from E3.5 and E4.5 uteri. **f, g**, Number (**f**) and location (**g**) of implantation sites at E4.5 and E5.5 uteri. Blue bands (arrows) indicate implantation sites; brackets indicate clustered implantation sites. Asterisk, $P < 0.001$. In all figures, error bars are standard deviations. +/+, +/- and -/- represent wild-type, heterozygote and LPA₃-deficient mice, respectively.

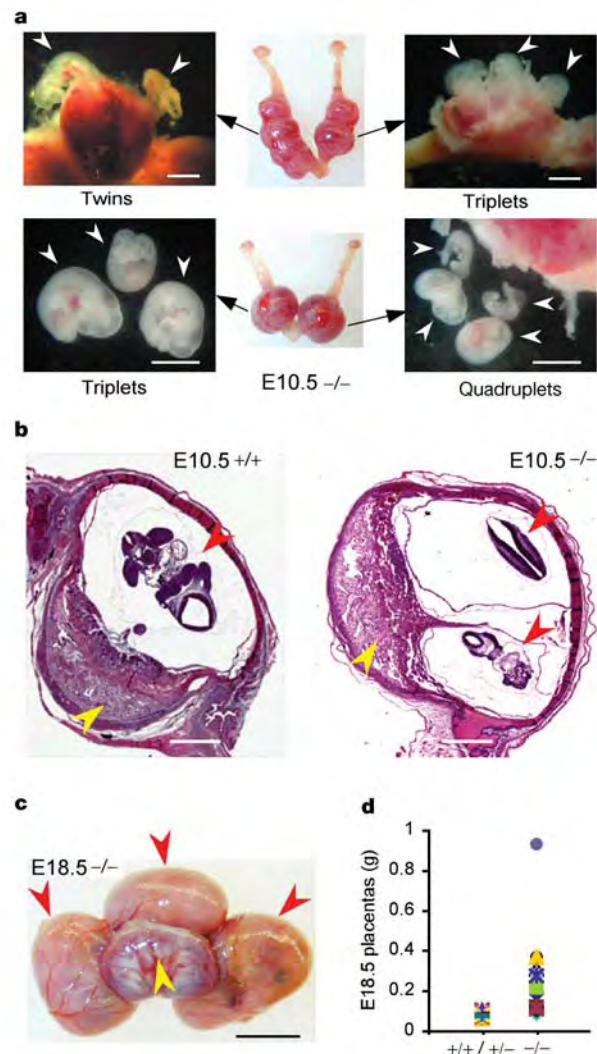


Figure 2 Multiple embryos at individual implantation sites and placental hypertrophy in uteri of LPA₃-deficient mice. **a**, Samples of multiple embryos at individual implantation sites at E10.5. White arrowheads indicate embryos. Scale bars: 2 mm. **b**, Cross-sections of E10.5 uteri revealing two less-developed embryos sharing one placenta in a LPA₃-deficient uterus. Scale bars: 1 mm. **c**, A placenta shared by three embryos at E18.5 in a LPA₃-deficient uterus. Scale bar: 8 mm. Red arrowheads in **b** and **c** indicate embryos; yellow arrowheads indicate placentas. **d**, Weight of placentas at E18.5.

three embryos (Fig. 2c) and associated with placental hypertrophy (Fig. 2d; see also Supplementary Fig. 5). These phenomena were never observed in controls.

In addition, embryos isolated from LPA₃-deficient uteri (at E10.5 and E18.5) were always smaller than those from wild-type or LPA₃ heterozygote controls at comparable ages (Fig. 3a–c), although newborns from LPA₃-deficient females were, on average, heavier (Fig. 3c), possibly resulting from prolonged pregnancy and/or smaller litter size. The average number of live embryos per animal that could be isolated from LPA₃-deficient females decreased after initial implantation (Fig. 3d). Delayed implantation and aberrant embryo spacing were thus associated with both delayed embryonic development and embryonic death, which could account for the reduced litter sizes produced by LPA₃-deficient females.

To determine whether LPA₃ loss in the embryo itself might contribute to the phenotypes, embryo transfer experiments were pursued. Wild-type blastocysts were transferred to wild-type or LPA₃-deficient pseudo-pregnant uteri. Transferred wild-type embryos in wild-type or LPA₃-deficient pseudo-pregnant uteri were phenotypically indistinguishable for implantation and development compared to embryos produced by natural matings in wild-type or LPA₃-deficient animals. Similarly, when LPA₃-deficient blastocysts were transferred to wild-type pseudo-pregnant uteri, no implantation or spacing abnormalities were observed (Supplementary Fig. 6 and data not shown). Wild-type and LPA₃-deficient blastocysts had comparable implantation rates in

wild-type pseudo-pregnant uteri. Considering the fact that no LPA₃ mRNA was detected in wild-type pre-implantation blastocysts, these results eliminate significant contributions of LPA₃ via pre-/post-implantation embryos, and indicate that maternal LPA₃ signalling is responsible for the observed phenotypes.

The observed implantation phenotypes of LPA₃-deficient mice were markedly similar to those reported for rats and mice treated with indomethacin^{17,18}, and for mice deficient in cytosolic phospholipase A_{2α} (cPLA_{2α})¹⁹. Indomethacin is an inhibitor of cyclooxygenases²⁰, which convert arachidonic acid to prostaglandin H₂ (PGH₂) in the biosynthesis of prostaglandins, whereas cPLA_{2α} is an important enzyme producing arachidonic acid. Notably, COX2 deficiency, but not COX1 deficiency, in mice results in multiple female reproductive failures, including implantation defects^{5,21}, although the precise phenotypes can be influenced by genetic background^{22,23}. Moreover, PGE₂ and carbaprostacyclin (cPGI, a stable analogue of PGI₂) can partially correct implantation defects in both cPLA_{2α}-deficient and COX2-deficient mice^{19,24}. These data indicate that the cPLA_{2α}–arachidonic acid–COX–prostaglandin pathway is crucial for implantation¹. In view of the phenotypic similarities between LPA₃ deficiency and cPLA_{2α}/prostaglandin

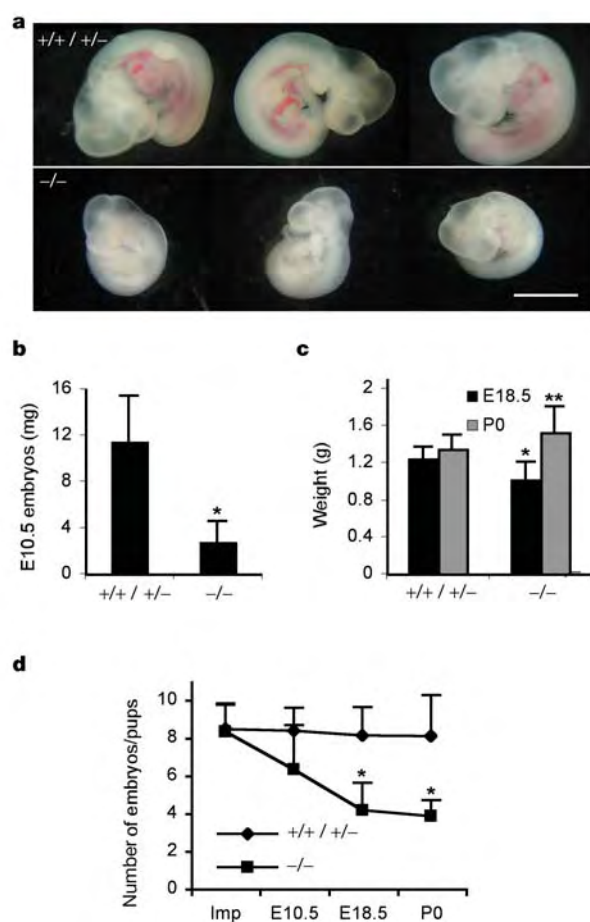


Figure 3 Delayed post-implantational development of embryos and increased embryonic death in uteri of LPA₃-deficient mice. **a**, Representative embryos from E10.5 uteri. Scale bar: 2 mm. **b**, E10.5 embryo weight. **c**, Weights of E18.5 embryos and postnatal day 0 (P0) pups. **d**, The average numbers of embryos implanted (Imp), at E10.5 and E18.5, and P0 pups. The numbers of embryos implanted were calculated as described in Methods. Asterisk, $P < 0.001$; double asterisk, $P < 0.05$.

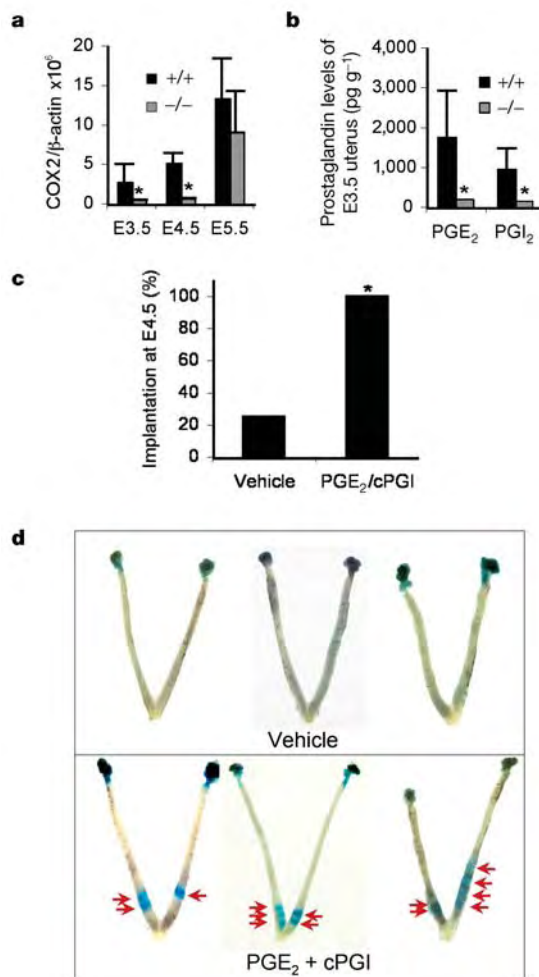


Figure 4 Reduced COX2 mRNA and prostaglandin levels in uteri of LPA₃-deficient mice, and exogenous prostaglandin rescue of delayed implantation. **a**, Expression of COX2 during early pregnancy in wild-type and LPA₃-deficient uteri. Asterisk, $P < 0.05$. **b**, Reduced PGE₂ and PGI₂ levels in E3.5 LPA₃-deficient uteri. Asterisk, $P < 0.05$. **c**, Significantly increased percentage of LPA₃-deficient female mice showing on-time implantation upon PGE₂ and cPGI (a stable PGI₂ analogue) treatment at E3.5. Asterisk, $P = 0.003$. **d**, Images of E4.5 LPA₃-deficient uteri with or without prostaglandin treatment. Red arrows indicate implantation sites.

deficiency, we proposed that LPA₃ might converge on this signalling pathway.

Components of prostaglandin signalling were therefore examined in uteri of LPA₃-deficient mice. These components included cPLA_{2α}, COX1 and COX2, and G-protein-coupled prostaglandin E₂ receptors EP_{1–4} and prostaglandin I₂ receptor, IP₂₅, along with leukaemia inhibitory factor (LIF) and Hoxa-10, two key regulators in implantation¹. Only COX2 mRNA levels were significantly reduced in LPA₃-deficient uteri (Fig. 4a, Supplementary Fig. 7, and data not shown). COX2 is a rate-limiting enzyme for prostaglandin biosynthesis. Consistent with this function, the suppression of COX2 expression in E3.5 LPA₃-deficient uteri resulted in reduced production of PGE₂ and PGI₂ (Fig. 4b), leading to conditions that are inadequate for implantation, which normally occurs around E4.0 (refs 18, 26). To rescue this prostaglandin reduction, we delivered exogenous PGE₂ and cPGI to E3.5 LPA₃-deficient female mice, a general approach previously reported¹⁹. After prostaglandin exposure, significantly more LPA₃-deficient female mice with normal (on-time) implantation were detected compared with LPA₃-deficient females given vehicle controls (Fig. 4c, $P = 0.003$). Notably, this rescue did not affect the uneven embryo spacing nor completely restore the reduction of implantation sites compared with wild-type controls (Figs 1g and 4d).

These findings identify LPA signalling as having an influence on embryo implantation, and are the first to link a lysophospholipid G-protein-coupled receptor to prostaglandin biosynthesis, thereby influencing female fertility. As a class, lysophospholipid receptors represent a 'drugable' target, as demonstrated by the compound FTY720, which is currently in phase III clinical trials for prevention of transplantation rejection²⁷. This raises the possibility of creating medicines that influence implantation timing, a critical factor for *in vitro* fertilization^{1,9} and also for reducing the incidence of multi-embryo gestations, especially monochorionic gestations, which can result in fetal demise⁶. The reduced litter sizes observed in receptor-null mutants for another lysophospholipid, sphingosine 1-phosphate, suggest that other lysophospholipid receptors may also influence mammalian reproduction through pharmacologically tractable mechanisms²⁸. □

Methods

Quantitative RT-PCR

Primers used were as described^{11,12,29}. For amplification of COX2, the following primers were used: forward 5'-AAGCGAGGACCTGGGTTC-3'; reverse, 5'-AAGCGCAGTTTATGTTGTCGT-3'. Quantitative RT-PCR was performed as described²⁹. The transcript number of target genes was quantified and normalized against GAPDH or β-actin transcript number.

In situ hybridization and histology

The animals were anaesthetized with halothane inhalation followed by cervical dislocation. *In situ* hybridization and histology were performed as described³⁰. Sense and antisense DIG-labelled cRNA probes were generated using appropriate polymerases from a full-length murine LPA₃ complementary DNA.

Mating, embryo collection and implantation localization

All the mice used in this study were of mixed background (129/Sv) and C57BL/6). Because no difference was observed for all the parameters examined between wild-type and heterozygote female mice (Supplementary Table 2, Supplementary Fig. 8a, b, and data not shown), females of either wild-type or heterozygote genotypes were used as controls. Females were naturally mated with wild-type stud males. The day a plug was found was designated as E0.5. Plugged females were anaesthetized with halothane inhalation followed by cervical dislocation. Uteri of pregnant females were dissected at E3.5, E4.5, E5.5, E10.5 and E18.5. Embryos at E10.5 were fixed in 10% formalin overnight before being weighed. Implantation sites at E4.5 and E5.5 were localized by intravenous injection of Evans blue dye (200 μl, 1% in 1 × PBS, Sigma)¹⁶. The number of embryos initially implanted in LPA₃-deficient and wild-type and heterozygote uteri were retrospectively calculated from E10.5 as follows: at E10.5, embryos in an average of 1.2 implantation sites (out of a total of 5.0) were absorbed in LPA₃-deficient uteri, but embryos in only 0.09 implantation sites (out of a total of 8.4) were absorbed in wild-type and heterozygote uteri ($P = 1.7 \times 10^{-5}$). With an average of 1.65 live embryos per implantation site in LPA₃-deficient uteri and 1.0 live embryo per implantation site in wild-type and heterozygote uteri at E10.5, the total number of embryos initially implanted should be 8.3 ((3.8

live + 1.2 absorbed) × 1.65) in LPA₃-deficient uteri and 8.4 ((8.31 live + 0.09 absorbed) × 1.0) in wild-type and heterozygote uteri.

Prostaglandin measurement

Uteri from E3.5 wild-type or LPA₃-deficient mice were immediately frozen and crushed in liquid nitrogen. Prostaglandins were extracted by the ethyl acetate extraction method. The prostaglandin levels of each sample were determined using the prostaglandin enzyme-linked immunoassay kit (Cayman Chemical). PGI₂ was measured as 6-keto-PGF_{1α}.

Prostaglandin administration

E3.5 LPA₃-deficient females were intraperitoneally injected with 100 μl of vehicle (10% ETOH with saline, as control) or 5 μg cPGI and 5 μg PGE₂ (Cayman Chemical, in 10% ETOH with saline) at 10:00 and 18:00. Implantation sites were detected using Evans blue dye at E4.5.

Data representation

Data are expressed as mean ± s.d. Statistical analyses were done using Student's *t*-test or χ^2 test. The significance level was set at $P < 0.05$.

Received 7 February; accepted 28 February 2005; doi:10.1038/nature03505.

- Dey, S. K. *et al.* Molecular cues to implantation. *Endocr. Rev.* **25**, 341–373 (2004).
- Ishii, I., Fukushima, N., Ye, X. & Chun, J. Lysophospholipid receptors: signaling and biology. *Annu. Rev. Biochem.* **73**, 321–354 (2004).
- Bandoh, K. *et al.* Molecular cloning and characterization of a novel human G-protein-coupled receptor, EDG7, for lysophosphatidic acid. *J. Biol. Chem.* **274**, 27776–27785 (1999).
- Contos, J. J. & Chun, J. The mouse lp(A3)/Edg7 lysophosphatidic acid receptor gene: genomic structure, chromosomal localization, and expression pattern. *Gene* **267**, 243–253 (2001).
- Lim, H. *et al.* Multiple female reproductive failures in cyclooxygenase 2-deficient mice. *Cell* **91**, 197–208 (1997).
- Cleary-Goldman, J. & D'Alton, M. Management of single fetal demise in a multiple gestation. *Obstet. Gynecol. Surv.* **59**, 285–298 (2004).
- Umstad, M. P. & Gronow, M. J. Multiple pregnancy: a modern epidemic? *Med. J. Aust.* **178**, 613–615 (2003).
- Daniel, Y. *et al.* Analysis of 104 twin pregnancies conceived with assisted reproductive technologies and 193 spontaneously conceived twin pregnancies. *Fertil. Steril.* **74**, 683–689 (2000).
- Wilcox, A. J., Baird, D. D. & Weinberg, C. R. Time of implantation of the conceptus and loss of pregnancy. *N. Engl. J. Med.* **340**, 1796–1799 (1999).
- Tokumura, A., Fukuzawa, K., Yamada, S. & Tsukatani, H. Stimulatory effect of lysophosphatidic acids on uterine smooth muscles of non-pregnant rats. *Arch. Int. Pharmacodyn. Ther.* **245**, 74–83 (1980).
- Contos, J. J., Fukushima, N., Weiner, J. A., Kaushal, D. & Chun, J. Requirement for the lpA1 lysophosphatidic acid receptor gene in normal suckling behavior. *Proc. Natl Acad. Sci. USA* **97**, 13384–13389 (2000).
- Contos, J. J. *et al.* Characterization of lpA₂ (Edg4) and lpA₁/lpA₂ (Edg2/Edg4) lysophosphatidic acid receptor knockout mice: signaling deficits without obvious phenotypic abnormality attributable to lpA₂. *Mol. Cell. Biol.* **22**, 6921–6929 (2002).
- Kingsbury, M. A., Rehen, S. K., Contos, J. J., Higgins, C. M. & Chun, J. Non-proliferative effects of lysophosphatidic acid enhance cortical growth and folding. *Nature Neurosci.* **6**, 1292–1299 (2003).
- Inoue, M. *et al.* Initiation of neuropathic pain requires lysophosphatidic acid receptor signaling. *Nature Med.* **10**, 712–718 (2004).
- Noguchi, K., Ishii, S. & Shimizu, T. Identification of p2y9/GPR23 as a novel G protein-coupled receptor for lysophosphatidic acid, structurally distant from the Edg family. *J. Biol. Chem.* **278**, 25600–25606 (2003).
- Paria, B. C., Huet-Hudson, Y. M. & Dey, S. K. Blastocyst's state of activity determines the "window" of implantation in the receptive mouse uterus. *Proc. Natl Acad. Sci. USA* **90**, 10159–10162 (1993).
- Kennedy, T. G. Evidence for a role for prostaglandins in the initiation of blastocyst implantation in the rat. *Biol. Reprod.* **16**, 286–291 (1977).
- Kinoshita, K. *et al.* Involvement of prostaglandins in implantation in the pregnant mouse. *Adv. Prostaglandin Thromboxane Leukot. Res.* **15**, 605–607 (1985).
- Song, H. *et al.* Cytosolic phospholipase A_{2α} is crucial for 'on-time' embryo implantation that directs subsequent development. *Development* **129**, 2879–2889 (2002).
- Frenkian, M., Segond, N., Pidoux, E., Cohen, R. & Jullienne, A. Indomethacin, a COX inhibitor, enhances 15-PGDH and decreases human tumoral C cells proliferation. *Prostaglandins* **65**, 11–20 (2001).
- Reese, J., Brown, N., Paria, B. C., Morrow, J. & Dey, S. K. COX-2 compensation in the uterus of COX-1 deficient mice during the pre-implantation period. *Mol. Cell. Endocrinol.* **150**, 23–31 (1999).
- Sibilia, M. & Wagner, E. F. Strain-dependent epithelial defects in mice lacking the EGF receptor. *Science* **269**, 234–238 (1995).
- Wang, H. *et al.* Rescue of female infertility from the loss of cyclooxygenase-2 by compensatory up-regulation of cyclooxygenase-1 is a function of genetic makeup. *J. Biol. Chem.* **279**, 10649–10658 (2004).
- Lim, H. *et al.* Cyclo-oxygenase-2-derived prostacyclin mediates embryo implantation in the mouse via PPAR δ . *Genes Dev.* **13**, 1561–1574 (1999).
- Yang, Z. M. *et al.* Potential sites of prostaglandin actions in the periimplantation mouse uterus: differential expression and regulation of prostaglandin receptor genes. *Biol. Reprod.* **56**, 368–379 (1997).
- Paria, B. C., Song, H. & Dey, S. K. Implantation: molecular basis of embryo-uterine dialogue. *Int. J. Dev. Biol.* **45**, 597–605 (2001).
- Gabardi, S. & Cerio, J. Future immunosuppressive agents in solid-organ transplantation. *Prog. Transplant.* **14**, 148–156 (2004).
- Ishii, I. *et al.* Marked perinatal lethality and cellular signaling deficits in mice null for the two sphingosine 1-phosphate (S1P) receptors, S1P₂/LPB₂/EDG-5 and S1P₃/LPB₃/EDG-3. *J. Biol. Chem.* **277**, 25152–25159 (2002).
- Hama, K. *et al.* Lysophosphatidic acid and autotaxin stimulate cell motility of neoplastic and non-neoplastic cells through LPA1. *J. Biol. Chem.* **279**, 17634–17639 (2004).

30. Ishii, I. *et al.* Selective loss of sphingosine 1-phosphate signaling with no obvious phenotypic abnormality in mice lacking its G protein-coupled receptor, LP₈₃/EDG-3. *J. Biol. Chem.* **276**, 33697–33704 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank F. Liu, S. Kupriyanov, R. Rivera, D. Herr, E. Nilsson, M. Murakami, Y. Kita, B. C. Paria, C. Akita, S. Carlson and Q. Chen for technical assistance and suggestions. This work was supported by grants from the National Institute of Mental Health to J.C. and J.J.A.C., the National Institute of Health to M.K.S, Swiss National Science Foundation to B.A., and Program for Promotion of Fundamental Studies in Health Sciences of the Pharmaceuticals and Medical Devices Agency (PMDA) and grants-in-aid from the Ministry of Education, Science, Culture and Sports for the 21st Century Center of Excellence Program, Japan, to H.S., J.A. and H.A.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C. (jchun@scripps.edu).

Cellular APOBEC3G restricts HIV-1 infection in resting CD4⁺ T cells

Ya-Lin Chiu¹, Vanessa B. Soros¹, Jason F. Kreisberg^{1,2}, Kim Stopak^{1,2}, Wes Yonemoto¹ & Warner C. Greene^{1,2}

¹Gladstone Institute of Virology and Immunology, ²Departments of Medicine, Microbiology and Immunology, University of California, San Francisco, California 94143, USA

In contrast to activated CD4⁺ T cells, resting human CD4⁺ T cells circulating in blood are highly resistant to infection with human immunodeficiency virus (HIV)^{1–4}. Whether the inability of HIV to infect these resting CD4⁺ T cells is due to the lack of a key factor, or alternatively reflects the presence of an efficient mechanism for defence against HIV, is not clear. Here we show that the anti-retroviral deoxycytidine deaminase APOBEC3G⁵ strongly protects unstimulated peripheral blood CD4⁺ T cells against HIV-1 infection. In activated CD4⁺ T cells, cytoplasmic APOBEC3G resides in an enzymatically inactive, high-molecular-mass (HMM) ribonucleoprotein complex that converts to an enzymatically active low-molecular-mass (LMM) form after treatment with RNase. In contrast, LMM APOBEC3G predominates in unstimulated CD4⁺ T cells, where HIV-1 replication is blocked and reverse transcription is impaired^{1–3}. Mitogen activation induces the recruitment of LMM APOBEC3G into the HMM complex, and this correlates with a sharp increase in permissivity for HIV infection in these stimulated cells. Notably, when APOBEC3G-specific small interfering RNAs are introduced into unstimulated CD4⁺ T cells, the early replication block encountered by HIV-1 is greatly relieved. Thus, LMM APOBEC3G functions as a potent post-entry restriction factor for HIV-1 in unstimulated CD4⁺ T cells. Surprisingly, sequencing of the reverse transcripts slowly formed in unstimulated CD4⁺ T cells reveals only low levels of dG → dA hypermutation, raising the possibility that the APOBEC3G-restricting activity may not be strictly dependent on deoxycytidine deamination.

APOBEC3G belongs to a family of tissue-restricted (deoxy)cytidine deaminases⁶ that edit RNA and mutates DNA^{6,7}. In the case of HIV, the incorporation of APOBEC3G into virions leads to extensive mutation of nascent HIV DNA formed during reverse transcription in the next round of infection^{5,8–11}. The Vif protein of HIV circumvents this anti-HIV defence mechanism by enhancing the 26S proteasome-mediated degradation of APOBEC3G^{12–14} and decreasing its synthesis^{12,15}. These events make APOBEC3G

unavailable for incorporation into budding virions, and thus, its anti-retroviral action is forfeited.

One puzzling aspect of APOBEC3G biology is why cytoplasmic forms of this enzyme in target CD4⁺ T cells undergoing HIV infection fail to recapitulate the antiviral effects of intra-virion APOBEC3G. Perhaps cellular APOBEC3G is subject to some form of negative regulation. APOBEC1 (ref. 16) and activation-induced cytidine deaminase (AID)¹⁷ are two well-characterized members of this (deoxy)cytidine deaminase family. APOBEC1 is the central component of an RNA-editing complex and is regulated by its assembly with APOBEC1-complementing factor¹⁶. AID catalyses the deamination of dC residues, yielding dU on single-stranded DNA *in vitro*, but has no measurable deaminase activity unless pre-treated with RNase to remove small inhibitory RNAs bound to AID¹⁸. These findings prompted us to determine whether APOBEC3G assembles into HMM complexes and to assess the potential role of RNA binding in the regulation of APOBEC3G.

Proteins in lysates of H9 T cells harbouring endogenous APOBEC3G were size fractionated by gel filtration on a Superose 6HR 10/30 column using a fast performance liquid chromatography (FPLC) apparatus followed by SDS–polyacrylamide gel electrophoresis (PAGE) and anti-APOBEC3G immunoblotting of the individual fractions. The endogenous ~46-kDa APOBEC3G enzyme resided principally in a HMM complex that is >700 kDa in mass (Fig. 1a, top). However, when lysates were treated with

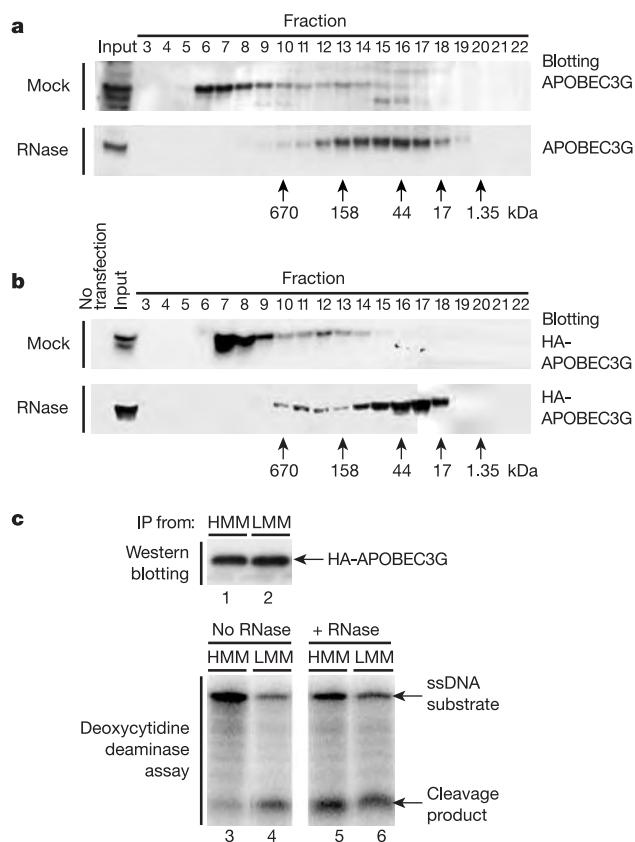


Figure 1 APOBEC3G is negatively regulated by recruitment into an enzymatically inactive HMM complex. **a**, **b**, Endogenous APOBEC3G in human H9 T cells (**a**) and exogenous HA-APOBEC3G expressed in 293T cells (**b**) reside principally in >700-kDa HMM complexes that can be converted to LMM forms after RNase A treatment. **c**, HMM (**b**, top, fractions 7–9) and LMM (**b**, bottom, fractions 15–17) forms of HA-APOBEC3G resolved by FPLC were immunoprecipitated (IP) with anti-HA (Covance), and HA-APOBEC3G protein content was assessed by anti-APOBEC3G antibody (lanes 1 and 2). The immunoprecipitates were tested in a deoxycytidine deaminase assay with or without RNase treatment. ssDNA, single-stranded DNA.

RNase A before fractionation, APOBEC3G was detected as a 46–100-kDa LMM form (Fig. 1a, bottom). This effect of RNase was blocked by the addition of a specific inhibitor of pancreatic-type RNase A (Supplementary Fig. 1). As shown by transfection of haemagglutinin-tagged APOBEC3G (HA-APOBEC3G) expression vector DNA¹², all of the putative cellular cofactors needed for the assembly of the HMM APOBEC3G complex appear to be expressed in 293T cells, which lack endogenous APOBEC3G expression (Fig. 1b).

Next, we examined the deoxycytidine deaminase activity of the two forms of APOBEC3G. In an *in vitro* assay, LMM APOBEC3G displayed enzymatic activity, but the HMM form did not (Fig. 1c, lanes 3 and 4). RNase treatment conferred activity on isolated HMM complexes but did not further increase the activity of the LMM form

(Fig. 1c, lanes 5 and 6). Thus, recruitment of APOBEC3G into the HMM ribonucleoprotein complex negatively regulates its enzymatic activity within cells, and degradation of the RNA component leads to disassembly of the complex and acquisition of enzymatic activity.

To assess the effect of Vif on the HMM APOBEC3G complex, we coexpressed HIV-1 Vif, HA-APOBEC3G and His₆-ubiquitin in 293T cells. In the presence of the proteasome inhibitor MG132, Vif co-fractionated with HMM APOBEC3G, and a ladder of potentially polyubiquitinated forms of the APOBEC3G enzyme was detected in the presence but not absence of Vif (Fig. 2a, middle and top panels). In the absence of APOBEC3G, Vif fractionated in a lower-molecular-mass range (Fig. 2a, bottom). Physical assembly of Vif and APOBEC3G in the HMM complex was confirmed by the

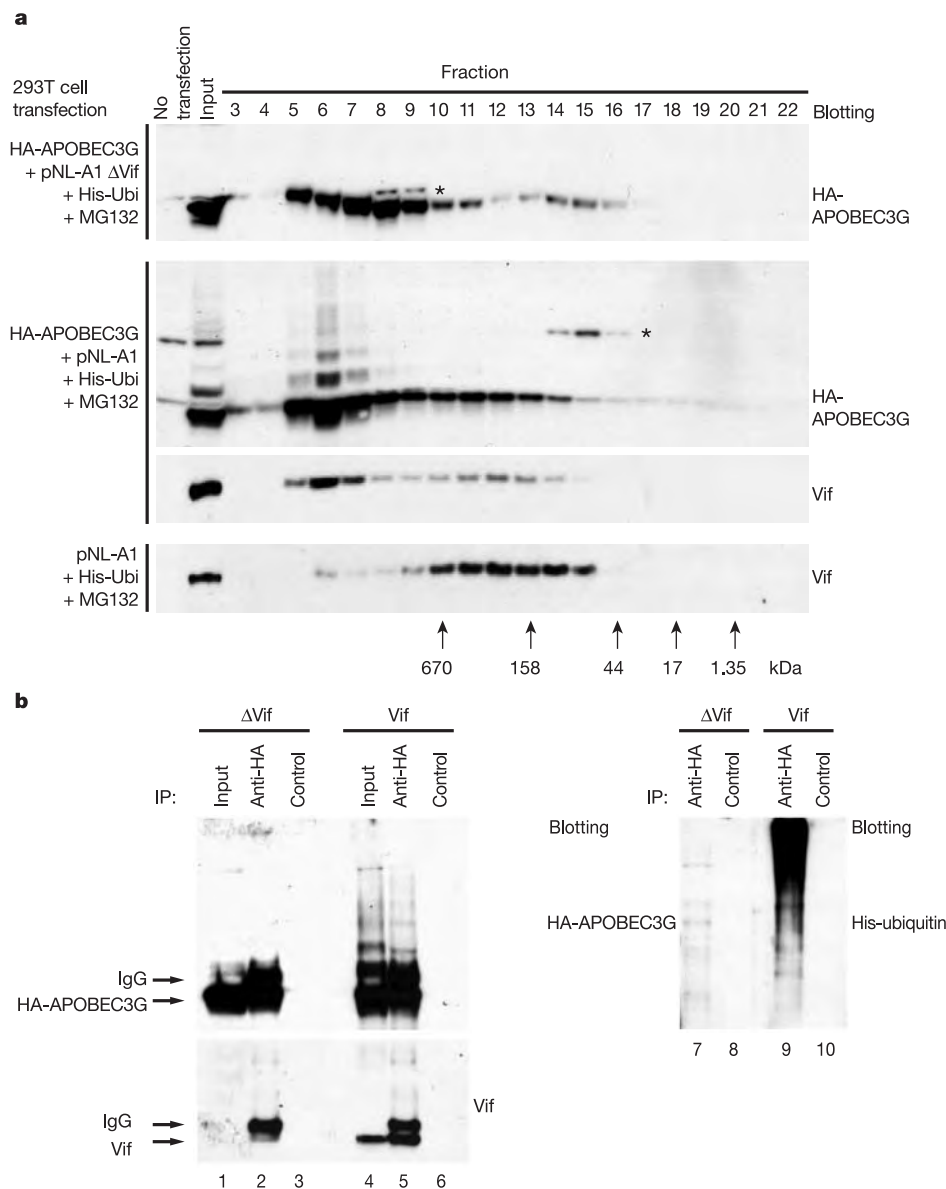


Figure 2 Vif assembles with APOBEC3G in HMM complexes and promotes polyubiquitination of APOBEC3G. **a**, Vif co-fractionates with the HMM APOBEC3G complex. 293T cells were transfected as indicated, cultured for 32 h, treated with proteasome inhibitor (MG132, 12.5 μ M) for 16 h, and subjected to FPLC analysis. Vif shifted to HMM fractions in the presence of HA-APOBEC3G, and a ladder of potentially polyubiquitinated forms of the APOBEC3G enzyme was detected (lane 6, upper middle panel). An asterisk indicates nonspecific bands detected by anti-HA antibody (12CA5, Roche). **b**, HMM fractions from 293T cells transfected with HA-APOBEC3G and His₆-

ubiquitin (His-Ubi) in the absence (pNL-A1 ΔVif) (**a**, top panel, lane 6) or presence of Vif (pNL-A1) (**a**, middle panels, lane 6) serve as input materials for an immunoprecipitation (IP) assay with anti-HA, resolved by SDS-PAGE, and immunoblotted with anti-HA, anti-Vif or anti-His antibodies (Santa Cruz Biotechnology). Protein G-agarose beads in the absence of anti-HA were used for control immunoprecipitation. In the left panel, note the co-immunoprecipitation of Vif and APOBEC3G (lane 5). In the right panel, when coexpressed with Vif, APOBEC3G in the HMM complex was polyubiquitinated (lane 9).

detection of Vif in anti-HA-APOBEC3G immunoprecipitates (Fig. 2b, lane 5). Additionally, polyubiquitinated forms of APOBEC3G in the HMM complex were detected when Vif and APOBEC3G were coexpressed (Fig. 2b, lane 9), but not when APOBEC3G was expressed in the absence of Vif (Fig. 2b, lane 7). We also demonstrated that Vif targets APOBEC3G for high-level polyubiquitination by recruitment of Cul5 and elongin B, which are components of an E3 ubiquitin ligase complex¹⁹, to the HMM complex (Supplementary Fig. 2).

We next surveyed primary cells for APOBEC3G complexes. Notably, only LMM forms of APOBEC3G were detected in purified, unstimulated CD4⁺ T cells and monocytes from peripheral blood (Fig. 3). However, when the CD4⁺ T cells were activated with phytohemagglutinin (PHA) and interleukin-2 (IL-2), RNase-sensitive HMM APOBEC3G complexes were detected (Fig. 3 and data not shown). When monocytes were induced to differentiate into macrophages by plastic adherence, HMM APOBEC3G complexes appeared (Fig. 3). These findings were of particular interest because activated CD4⁺ T cells and macrophages readily support HIV replication, whereas unstimulated peripheral blood T cells and freshly isolated monocytes do not. This phenomenon is due, at least in part, to an early post-entry block occurring at or soon after the reverse transcription step^{1–4,20}.

To explore further the effects of CD4⁺ T-cell activation on HMM APOBEC3G complex formation and single-cycle HIV infectivity, we used vesicular stomatitis virus-G (VSV-G)-pseudotyped NL4-3 viruses expressing a heat-stable antigen (HSA, mouse CD24) reporter gene in lieu of Nef (Supplementary Fig. 3). Co-stimulation of CD4⁺ T cells with anti-CD3 and anti-CD28 antibodies strongly induced HMM APOBEC3G complex formation, and these cells were highly permissive for HIV infection (Fig. 3; see also Supplementary Fig. 3). Anti-CD3 antibodies modestly induced HMM

complex assembly; anti-CD28 antibodies did not (Fig. 3). When cells were pre-treated with aphidicolin to block cell-cycle progression at the G1/S border or *n*-acetyl butyric acid (*n*-butyrate) to block progression from G1a to G1b (ref. 2) before anti-CD3 and anti-CD28 co-stimulation, markedly different results were obtained. Aphidicolin-treated cells displayed HMM APOBEC3G complexes and were permissive for HIV infection, whereas *n*-butyrate-treated cells contained only LMM forms of APOBEC3G and were not permissive (Fig. 3; see also Supplementary Fig. 3). These results suggest that progression of cells into G1b is necessary for HMM APOBEC3G complex formation and, as previously reported, HIV infection², but entry into S phase is not required for either event. Activation of CD4⁺ T cells with phorbol 12-myristate 13-acetate also induced HMM APOBEC3G complex formation (Fig. 3) and increased permissivity for HIV infection (Supplementary Fig. 3).

These composite findings raised the possibility that cellular LMM APOBEC3G functions as an effective early post-entry restriction factor for HIV in unstimulated peripheral blood CD4⁺ T cells and freshly isolated monocytes. To test this hypothesis, we introduced APOBEC3G-specific small interfering RNA (siRNA; APOBEC3G-si240) into unstimulated CD4⁺ T cells using an AMAXA nucleofactor apparatus. In studies with fluorescein isothiocyanate (FITC)-conjugated siRNA, about 80% of the nucleofected cells acquired the siRNA (data not shown). siRNA targeting APOBEC3G decreased APOBEC3G protein expression by a maximum of 70% between 48 and 88 h after nucleofection (Fig. 4a, b). Despite equivalent cellular uptake, a mutant siRNA (APOBEC3G-si240 mm) differing by two ribonucleotides did not alter APOBEC3G expression (Fig. 4a). Thus, the expression of endogenous APOBEC3G was decreased by nearly 90% in the cells acquiring the APOBEC3G-specific siRNA. Importantly, nucleofection did not activate the CD4⁺ T cells as

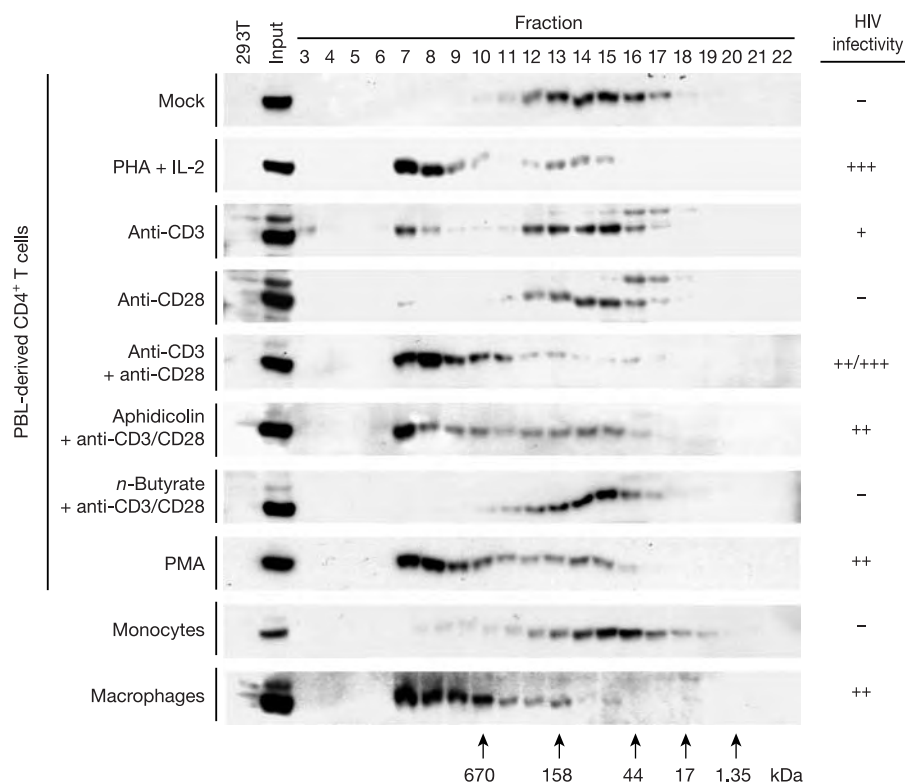


Figure 3 Inducibility of HMM APOBEC3G complex formation in primary CD4⁺ T cells with different stimuli, and correlation of the HMM complex with permissivity for HIV infection. APOBEC3G complexes were identified by FPLC, SDS-PAGE and immunoblotting with anti-APOBEC3G antibody. Single-round infection experiments were performed with the VSV-G-pseudotyped NL4-3 HSA virus. Ranges of HIV infectivity were determined by

normalizing the percentage of HSA-positive cells obtained with each treatment to the response in cells stimulated with PHA and IL-2 (assigned a value of 100%). HIV infectivity: >70% (high, +++); 30–70% (moderate, ++); 5–30% (low, +); <5% (none, –). Scoring was based on the infectivity data obtained from cells isolated from four healthy donors. PMA, phorbol 12-myristate 13-acetate.

assessed by either anti-CD25 or anti-CD69 immunostaining (Fig. 4c).

To assess the susceptibility of unstimulated peripheral blood CD4⁺ T cells to HIV infection after siRNA treatment, we performed single-round infection experiments with the VSV-G-pseudotyped NL4-3 HSA reporter virus. As expected, unstimulated CD4⁺ T cells did not support reporter virus expression, but cells activated with PHA and IL-2 were readily infected (Fig. 4c). Of note, APOBEC3G-specific, but not mutant, siRNA markedly relieved the HIV replication block in these unstimulated CD4⁺ T cells (Fig. 4c). As expected, APOBEC3G siRNA-dependent expression of HSA required effective reverse transcription of the reporter virus because surface HSA

expression did not occur in cells previously treated with 3TC, a reverse transcriptase inhibitor (Fig. 4c).

Next, studies were performed to pinpoint the replication block induced by LMM APOBEC3G in unstimulated peripheral blood CD4⁺ T cells. Studies of HIV virion binding and fusion in cells treated with the mutant versus APOBEC3G-specific siRNA revealed no differences (data not shown). These results indicate that the APOBEC3G siRNA relieves a block occurring after virion binding and fusion in the unstimulated CD4⁺ T cells. Taqman-based real-time polymerase chain reaction (PCR) was used to quantify the synthesis of HIV-1 complementary DNA in cells nucleofected with APOBEC3G-specific or control siRNA. To limit infection to a

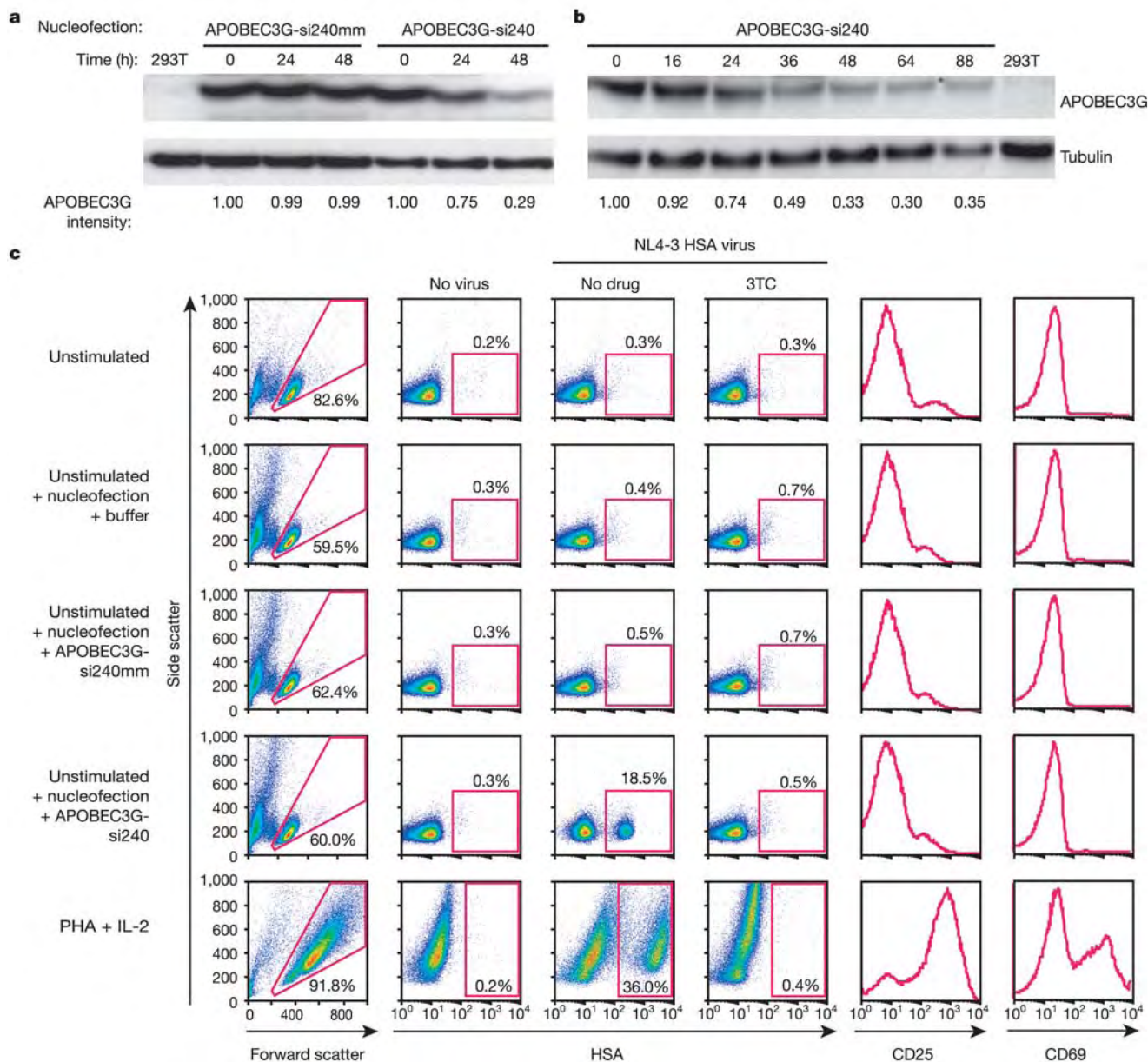


Figure 4 APOBEC3G is a post-entry restriction factor for HIV in unstimulated CD4⁺ T cells. **a**, APOBEC3G protein levels (normalized to tubulin) in lysates of PBL-derived CD4⁺ T cells nucleofected with APOBEC3G-specific (APOBEC3G-si240) or mutant siRNAs (APOBEC3G-si240mm; specificity control). **b**, Time course of siRNA-mediated knockdown of APOBEC3G in PBL-derived CD4⁺ T cells. **c**, siRNA-mediated knockdown of APOBEC3G renders unstimulated CD4⁺ T cells permissive for HIV infection. PBL-derived CD4⁺ T cells nucleofected with buffer or siRNAs were cultured in the absence or presence

of the reverse transcriptase inhibitor 3TC (10 μ M) and infected with NL4-3 HSA reporter virus 48 h after siRNA treatment. Untreated cells not receiving siRNA (unstimulated) served as a negative control, and cells stimulated with PHA and IL-2 as a positive control. Expression of the HSA reporter gene and the activation status of CD4⁺ T cells were analysed 48 h after infection. Similar results were obtained using cells from four different donors.

siRNA (data not shown). These results suggest the presence of a late reverse transcription defect in unstimulated CD4⁺ T cells that is alleviated upon siRNA-mediated knockdown of endogenous APOBEC3G expression. Delayed accumulation of late reverse transcription products has also been described in monocytes²⁰, strengthening the possibility that LMM APOBEC3G also mediates the replication delay in these cells.

a

Late RT product

Number of copies per 5x10⁵ cells

Time (h)

● PHA + IL-2
○ Nucleofection + APOBEC3G-si240
▲ Nucleofection + APOBEC3G-si240mm
◇ Nucleofection + buffer
■ Unstimulated

b

6581

Reference AAATTAACCCCACTCTGTGTAGTTTAAAGTGCACTGATTGAAGAATGATACTAATACCAATAGTAGTAGCGGGAGAATGATAATGGAGA

225TA-07A.....A...A..A..A.....AAA.A.....AA.A.

225TA-35A.....A...A..A..A.....AAA.A.....AA.A.

225TA-39A.A..A...A..A..A..A.....G.....A.AAA.A..A....AA.A.

225TA-40A.....A...A..A..A.....AAA.A.....AA.A.

348TA-31A.....A..A..A..A.....AAA.A..A....AA.A.

348TA-32A.....A...A..A..A.....A.....AAA.A..A....AA.A.

351TA-34G.A.....AAA.A..A....AA.A.

351TA-39A.....A..A..A..A.....AAA.A..A....AA.A.

351TA-40 ..G.....A.....A...A..A..A.....AAA.A..A....AA.A.

6672

Reference AAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAAGCATAAAGGATAGAGTGCAGAAAGAATATGCATTCTTTTATAAACCTTGATAT

225TA-07 ...AA.A.A.....A.A..A..A..A.....A.

225TA-35 ...AA.A.A.....A.A..A..A..A.....A.

225TA-39 ...AA.A.A.....A...A..A..A..A.....A.

225TA-40A.....A..A.....A.

348TA-31 ...AA.A.A.....A.A..A..A..A.....A.

348TA-32 ...AA.A.A.....A.A..A..A..A.....A.

351TA-34T.....A.A.....A.

351TA-39 ...AA.A.A.....A.A..A..A..A.....A.

351TA-40 ...AA.A.A.....A.A..A..A..A.....A.

6763

Reference AGTACCAATAGATAATAACAGCTATAGGTTGATAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAGGATATCCT

225TA-07A.....A..A.....AA.....

225TA-35A.....A..A.....AA.....

225TA-39A.....AA..A..A.....A.....AA.....

225TA-40A.....A..A.....A.....

348TA-31A..A.....A.....A.....

348TA-32A...G.....A..A.....A.....A.....

351TA-34A.....A.....A.....

351TA-39A.....A..A.....AA.....

351TA-40A.....A..A..A.....A.....

represent the means $\pm$ s.d. derived from assays independently performed three times.

b, Sequence analysis of the *env* region (6581–6842) of viral reverse transcripts synthesized in unstimulated PBL-derived CD4⁺ T cells of three donors. Nine of 110 (8%) independent *env* sequences displayed extensive dG $\rightarrow$ dA hypermutations. Sequences are compared to the starting NL4-3 HSA plasmid, and only nucleotides differing from this sequence are shown.

The LMM form of APOBEC3G functions as a highly active post-entry restriction factor for HIV in unstimulated peripheral blood CD4⁺ T cells and probably in freshly isolated monocytes. Interestingly, unstimulated CD4⁺ T cells present in human lymphoid tissues are permissive for HIV infection²¹, and APOBEC3G is predominantly in the HMM form in these cells (J. F. Kreisberg *et al.*, unpublished data). The cause of these differences is currently under investigation. The detection of dG → dA hypermutations in only 8% of the late reverse transcription products suggests that the replication block induced by cellular LMM APOBEC3G involves a non-enzymatic mechanism. Indeed, APOBEC3G-mediated DNA editing may not always be necessary for its antiviral activity²². One strong possibility involves the RNA binding properties of APOBEC3G. Binding of LMM APOBEC3G to the viral RNA in the reverse transcription complex could decrease the efficiency of reverse transcriptase action, leading to a kinetic delay in the accumulation of late reverse transcription products. This model could also explain why subsequent mitogen activation of infected resting CD4⁺ T cells rescues productive viral infection¹⁻³. Under these conditions, APOBEC3G may be recruited into the HMM complex, thus relieving the late block in reverse transcription. Viral inhibition mediated through the RNA-binding properties of APOBEC3G has also been described in the setting of the hepatitis B virus²³. Why the accumulating reverse transcripts in unstimulated CD4⁺ T cells display such unexpectedly low levels of dG → dA hypermutation is unclear. It is possible that the effectiveness of deoxycytidine deamination differs depending on whether the enzyme is originally recruited into virions or resides in the cytoplasm. Alternatively, our studies would underestimate the number of late reverse transcription products containing dG → dA hypermutations if these deaminated minus-strand precursors are particularly sensitive to degradation by uracil N-glycosylase and apurinic-apyrimidinic endonuclease in unstimulated CD4⁺ T cells. Such an effect would be consistent with the accelerated decay of HIV reverse transcripts previously described¹⁻³. APOBEC3F is a deoxycytidine deaminase that also exerts antiviral activity, albeit at levels lower than APOBEC3G²⁴⁻²⁷. APOBEC3G and APOBEC3F are coordinately expressed in a wide range of human tissues, including unstimulated T cells, and APOBEC3F-signature hypermutations are detectable in the viral cDNA sequences isolated from unstimulated CD4⁺ T cells (Fig. 5b). Thus, it will be important to explore whether APOBEC3F is also negatively regulated within a latent HMM complex and whether it similarly functions as a post-entry restriction factor for HIV in unstimulated CD4⁺ T cells and monocytes.

action of LMM APOBEC3G effectively impairs the growth of both wild-type and Δ Vif forms of HIV. Our findings also raise an intriguing possibility: agents that promote the disassembly of the HMM APOBEC3G complex might confer resistance to HIV infection in cells normally highly permissive for HIV infection. However, it will be important to exclude an effect of such disassembled forms of APOBEC3G on host genomic DNA in these activated and often proliferating cellular hosts. $\square$

A HIV-1-based reporter virus (NL4-3 HSA R-E-) encoding HSA (mouse CD24) was pseudotyped with the VSV-G envelope by calcium phosphate-mediated co-transfection of 293T cells with expression DNA plasmids. After 48 h, virus-containing supernatants were collected, clarified by sedimentation, filtered and concentrated by ultra-centrifugation. The viral titre was measured by anti-p24^{Gag} ELISA (PerkinElmer). Twenty-four hours after nucleofection with siRNA, CD4⁺ T cells were cultured with or without reverse transcription inhibitor (3TC, 10 μ M) for 24 h, incubated with reporter virus (200 ng of p24 per 0.5×10^6 cells) for 3 h, washed extensively, and cultured for 48 h. Unstimulated cells served as negative controls, and cells activated with PHA and IL-2 as positive controls. Expression of the reporter gene product and the activation status of CD4⁺ T cells were monitored by flow cytometry using FITC-conjugated anti-HSA (Pharmingen), phycoerythrin-conjugated anti-CD25 and allophycocyanin-conjugated anti-CD69 (BD Biosciences), respectively. Data were collected on a FACS Calibur (BD Biosciences) and analysed using Flowjo software (Treestar).

Real-time PCR quantification of reverse transcripts

NL4-3 HSA viral stocks prepared by transfection of 293T cells were treated with 60 U ml⁻¹ of Turbo DNase I (Ambion) for 1 h at 37 °C to remove residual plasmid DNA. Cells (0.5 × 10⁶) were infected as described above, and total DNA was prepared at the indicated time points and stored in 100 µl of 1 mM Tris, pH 7.4. Early and late HIV-1 reverse transcription products were quantified using specific primer and probe combinations. Reaction mixtures (25 µl) contained QuantiTect Probe PCR master mix (Qiagen), 300 nM primers, 200 nM probe and 5 µl of total DNA. PCR was performed for 15 min at 95 °C followed by 40 cycles of 15 s at 95 °C, and 1 min at 60 °C using an ABI Prism 7700 (Applied Biosystems). The specific amplification of newly synthesized reverse transcripts was monitored by treatment with 3TC.

DNA sequencing

DNA was obtained as for Taqman PCR and subjected to amplification (AccuPrime Taq, Invitrogen) with primers specific for the *env* region, cloned into pCRII-TOPO TA vector, and sequenced (Molecular Cloning Laboratory).

Received 9 November 2004; accepted 21 February 2005; doi:10.1038/nature03493.
Published online 13 April 2005.

1. Zack, J. A. *et al.* HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell* **61**, 213–222 (1990).
2. Korin, Y. D. & Zack, J. A. Progression to the G1b phase of the cell cycle is required for completion of human immunodeficiency virus type 1 reverse transcription in T cells. *J. Virol.* **72**, 3161–3168 (1998).
3. Pierson, T. C. *et al.* Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection. *J. Virol.* **76**, 8518–8531 (2002).
4. Stevenson, M., Stanwick, T. L., Dempsey, M. P. & Lamont, C. A. HIV-1 replication is controlled at the level of T cell activation and proviral integration. *EMBO J* **9**, 1551–1560 (1990).
5. Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650 (2002).
6. Jarmuz, A. *et al.* An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285–296 (2002).
7. Harris, R. S., Petersen-Mahrt, S. K. & Neuberger, M. S. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* **10**, 1247–1253 (2002).
8. Harris, R. S. *et al.* DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803–809 (2003).
9. Mangeat, B. *et al.* Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99–103 (2003).
10. Lecossier, D., Bouchonnet, F., Clavel, F. & Hance, A. J. Hypermutation of HIV-1 DNA in the absence of the Vif protein. *Science* **300**, 1112 (2003).
11. Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98 (2003).
12. Stopak, K., de Noronha, C., Yonemoto, W. & Greene, W. C. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Mol. Cell* **12**, 591–601 (2003).
13. Marin, M., Rose, K. M., Kozak, S. L. & Kabat, D. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nature Med.* **9**, 1398–1403 (2003).
14. Sheehy, A. M., Gaddis, N. C. & Malim, M. H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature Med.* **9**, 1404–1407 (2003).
15. Mariani, R. *et al.* Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* **114**, 21–31 (2003).
16. Anant, S. & Davidson, N. O. Molecular mechanisms of apolipoprotein B mRNA editing. *Curr. Opin. Lipidol.* **12**, 159–165 (2001).
17. Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
18. Branstetter, R., Pham, P., Scharff, M. D. & Goodman, M. E. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl Acad. Sci. USA* **100**, 4102–4107 (2003).
19. Yu, X. *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056–1060 (2003).
20. Triques, K. & Stevenson, M. Characterization of restrictions to human immunodeficiency virus type 1 infection of monocytes. *J. Virol.* **78**, 5523–5527 (2004).
21. Eckstein, D. A. *et al.* HIV-1 actively replicates in naive CD4(+) T cells residing within human lymphoid tissues. *Immunity* **15**, 671–682 (2001).
22. Newman, E. N. *et al.* Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr. Biol.* **15**, 166–170 (2005).
23. Turelli, P., Mangeat, B., Jost, S., Vianin, S. & Trono, D. Inhibition of hepatitis B virus replication by APOBEC3G. *Science* **303**, 1829 (2004).
24. Bishop, K. N. *et al.* Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.* **14**, 1392–1396 (2004).
25. Liddament, M. T., Brown, W. L., Schumacher, A. J. & Harris, R. S. APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 *in vivo*. *Curr. Biol.* **14**, 1385–1391 (2004).
26. Wiegand, H. L., Doehe, B. P., Bogerd, H. P. & Cullen, B. R. A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J.* **23**, 2451–2458 (2004).
27. Zheng, Y. H. *et al.* Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J. Virol.* **78**, 6073–6076 (2004).
28. Stremlau, M. *et al.* The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848–853 (2004).
29. Ganesh, L. *et al.* The gene product Murr1 restricts HIV-1 replication in resting CD4⁺ lymphocytes. *Nature* **426**, 853–857 (2003).
30. Camaur, D. & Trono, D. Characterization of human immunodeficiency virus type 1 Vif particle incorporation. *J. Virol.* **70**, 6106–6111 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank K. Strebel for the gift of the pNL-A1 plasmid, X.-F. Yu for the gift of the Myc-Cul5 plasmid, D. Bohmann for the gift of the His₆-tagged ubiquitin expression vector DNA, N. R. Landau for the gift of NL4-3 HSA R-E- reporter provirus through the AIDS Research and Reference Reagent Program at NIH, J. Burns for the gift of expression plasmid pVSV-G, and M. Malim for providing the monoclonal anti-Vif (319) antibody. We thank A. O'Mahony and H. Kwon for assistance with the FPLC experiments; J. Neidleman for assistance preparing the primary cells; M. Cavois, D. Fenard, A. Yonezawa, J. Bohuslav, L.-F. Chen, C. Martin and S. Williams for discussions; G. Howard and S. Ordway for editorial assistance; and S. Cammack, R. Givens and J. Carroll for assistance in preparation of the manuscript and the graphics. Different components of this work were supported by funding from the National Institutes of Health (Women's HIV Interdisciplinary Network and NIMH; W.C.G.) and the Universitywide AIDS Research Program (W.C.G.) and the American Foundation for AIDS Research (Y.-L.C.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to W.C.G. (wgreene@gladstone.ucsf.edu).

The bipolar mitotic kinesin Eg5 moves on both microtubules that it crosslinks

Lukas C. Kapitein¹, Erwin J. G. Peterman¹, Benjamin H. Kwok², Jeffrey H. Kim², Tarun M. Kapoor^{2*} & Christoph F. Schmidt^{1*}

¹Department of Physics and Astronomy and Laser Centre, Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

²Laboratory of Chemistry and Cell Biology, The Rockefeller University, New York, New York 10021, USA

* These authors contributed equally to this work

During cell division, mitotic spindles are assembled by microtubule-based motor proteins^{1,2}. The bipolar organization of spindles is essential for proper segregation of chromosomes, and requires plus-end-directed homotetrameric motor proteins of the widely conserved kinesin-5 (BimC) family³. Hypotheses for bipolar spindle formation include the 'push-pull mitotic muscle' model, in which kinesin-5 and opposing motor proteins act between overlapping microtubules^{2,4,5}. However, the precise roles of kinesin-5 during this process are unknown. Here we show that the vertebrate kinesin-5 Eg5 drives the sliding of microtubules depending on their relative orientation. We found in controlled *in vitro* assays that Eg5 has the remarkable capability of simultaneously moving at ~20 nm s⁻¹ towards the plus-ends of each of the two microtubules it crosslinks. For anti-parallel microtubules, this results in relative sliding at ~40 nm s⁻¹, comparable to spindle pole separation rates *in vivo*⁶. Furthermore, we found that Eg5 can tether microtubule plus-ends, suggesting an additional microtubule-binding mode for Eg5. Our results demonstrate how members of the kinesin-5 family are likely to function in mitosis, pushing apart interpolar microtubules as well as recruiting microtubules into bundles that are subsequently polarized by relative sliding.

Unlike most other kinesins characterized thus far, kinesin-5 proteins have four identical motor domains configured like a 'dumb-bell', with two motor domains at each end of a rod⁷ (Fig. 1a). Based on this bipolar structure, it has long been suggested that kinesin-5 proteins contribute to the bipolar organization of mitotic spindles by crosslinking and sliding overlapping interpolar microtubules^{7,8}. Evidence for kinesin-5-induced crosslinking comes from electron microscopy⁸. 'Whole system' approaches using cell extracts and pharmacological inhibition suggest an essential role for Eg5 in spindle morphogenesis and poleward flux^{9–14}. Such

approaches, however, do not directly explore how the motors work. It remains unknown whether kinesin-5 motors act as individual molecules between microtubules, or in some aggregated form, or whether they are associated with other structures¹². These questions need to be resolved by well-controlled *in vitro* assays.

To directly test the sliding hypothesis, we have set up an *in vitro* assay with microtubules and purified Eg5, a vertebrate kinesin-5. In the assay, axonemes (bundles of microtubules) are attached to a glass surface, and motors and fluorescent microtubules bind to them from solution (Fig. 1a). Fixing one of the interacting partners (in this instance, the axoneme) to the surface as a track for the other eliminates translational diffusion and allows motor-driven movements to be resolved clearly. One of the main challenges in developing the assay was to make sure that motors and microtubules bind only to fixed axonemes and not to the glass surface. This problem was solved by blocking the surface with a

polyethylene-glycol polymer brush after axoneme attachment.

We found that single microtubules readily bound to and aligned with axonemes in the presence of Eg5. Approximately half of the attached microtubules were immobile or moved very slowly ($<10 \text{ nm s}^{-1}$). The rest, roughly equal in number to the immobile ones, moved along the axonemes with an average speed of $40.2 \pm 1.8 \text{ nm s}^{-1}$ (mean $\pm$ s.e.m., $n = 52$) (Figs 1b and 2a). To exclude artefacts generated by the use of axonemes (for example, dynein contamination), we performed two controls. First, the same assay conditions without Eg5 resulted in no microtubule binding or sliding. Second, we observed relative sliding between microtubules in samples containing no axonemes (Fig. 1d, f). To exclude the possibility that aggregates of Eg5 produced the observed cross-linking and sliding, we performed two more controls. First, we used Eg5 immediately after gel filtration, taking only the mono-dispersed peak fraction of the elution profile¹⁴. This control yielded results consistent with those found with frozen-and-thawed motors. Second, we repeated the microtubule–microtubule interaction experiment using a truncated form of Eg5 (amino acids 1–591) (ref. 15) that on the basis of its sequence is not likely to form a homotetramer, and found no crosslinking of microtubules in the presence of ATP. Motility in surface gliding assays was comparable to that of full-length motors (data not shown). Together, our results demonstrate that full-length, tetrameric, bipolar Eg5 can crosslink microtubules, align them and drive their relative sliding.

To understand Eg5 function in spindle assembly, one must know how the relative orientation of microtubules affects bundling and sliding. It is tempting to hypothesize that a bipolar motor with catalytic domains at two opposite ends generates force and motion independently at both ends. We therefore looked for motility events in geometries where the motions along the two tracks could be separated. This was possible in the cases where microtubules were not aligned with the axis of the microtubule they were linked to. The two velocities of the crossing point with respect to both filaments could then be separately measured, provided that this point had not yet reached the end of one track. The average speed along the respective microtubules in such cases was $20.0 \pm 1.0 \text{ nm s}^{-1}$ (mean $\pm$ s.e.m., $n = 40$) (Fig. 2a–c), with no apparent dependence on the relative angle between the tracks. This observation shows that Eg5 can simultaneously move on both of the microtubules it crosslinks, with a speed comparable to that observed in surface gliding assays ($24.0 \pm 0.4 \text{ nm s}^{-1}$, $n = 46$) (arrow in Fig. 2c). In the case of two aligned filaments, one would expect a relative sliding velocity corresponding to either the sum (anti-parallel) or the difference (parallel) of the single-end speeds. This is consistent with the bimodal distribution ($<10 \text{ nm s}^{-1}$ and $\sim 40 \text{ nm s}^{-1}$) of relative velocities we observed in the aligned cases (Figs 1b and 2c). To probe directly the correlation between relative orientation and speed, we performed assays using polarity-marked microtubules (Fig. 1c). These experiments confirmed that parallel microtubules remained static, whereas an anti-parallel configuration was required for double-speed relative gliding (Fig. 1d–f).

To observe transitions in sliding velocity for a given pair of microtubules, we looked for changes in microtubule–microtubule orientation. This happened spontaneously in several cases when the moving microtubule flipped by 180° , usually when it had moved beyond the end of the axoneme. In these cases microtubules switched from moving to being stationary, which is consistent with changing from anti-parallel to parallel orientation (see Supplementary Video S1). In order to have direct control over the relative orientations between microtubules, we used optical tweezers to manipulate microtubules through attached handle beads. Figure 2d shows a twofold decrease in the velocity along the microtubule axis when the relative orientation changed from anti-parallel to orthogonal. In another case, we were able to start motility by inverting the orientation of a microtubule that had been immobile near the end of an axoneme. We cannot derive probabil-

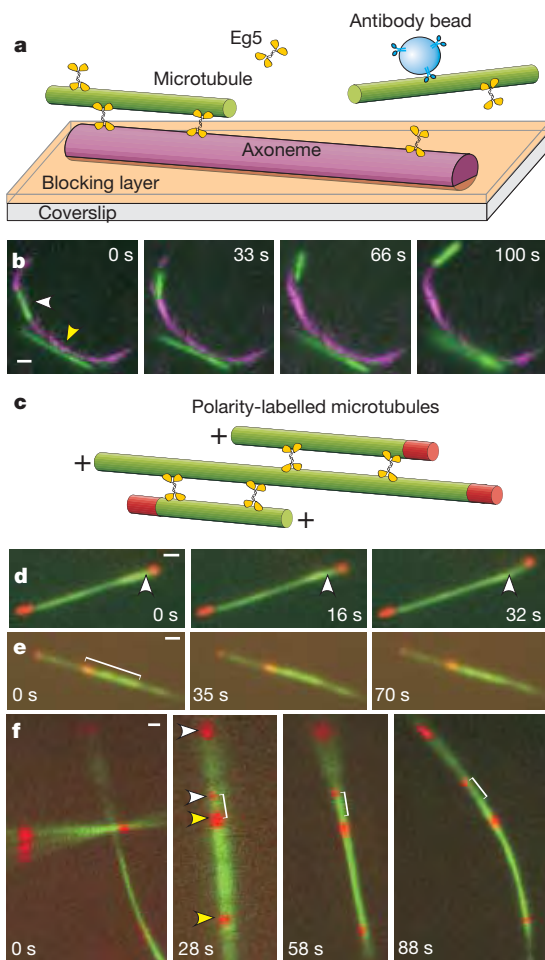


Figure 1 Eg5 can slide microtubules apart. **a**, Sketch of the *in vitro* assay with microtubules (green) attached via Eg5 motors (yellow) to surface-immobilized axonemes (magenta). The coverslip surface is blocked using a polymer brush. Beads (1- μm diameter, blue) coated with anti-tubulin antibodies were used in some experiments for manipulation with optical tweezers. **b**, Time-lapse images of both a sliding (white arrow, 40 nm s^{-1}) and a static (yellow arrow) fluorescent microtubule on a darkfield-detected axoneme. **c**, Sketch of the *in vitro* assay with polarity-marked microtubules. **d**, Anti-parallel microtubules sliding apart. The arrow marks the plus-end of the long microtubule, relative to which the short one moved at 35 nm s^{-1} . **e**, Two parallel microtubules (one marked with a white line) that were crosslinked and remained static. **f**, Time course of sliding within a bundle of microtubules. Two bundles first joined and aligned (left panel). Seeds marked with arrows of the same colour remained stationary relative to each other, but moved at 36 nm s^{-1} relative to those marked with a different colour. Scale bars, $1 \mu\text{m}$. See also Supplementary Videos S1, S2 and S3.

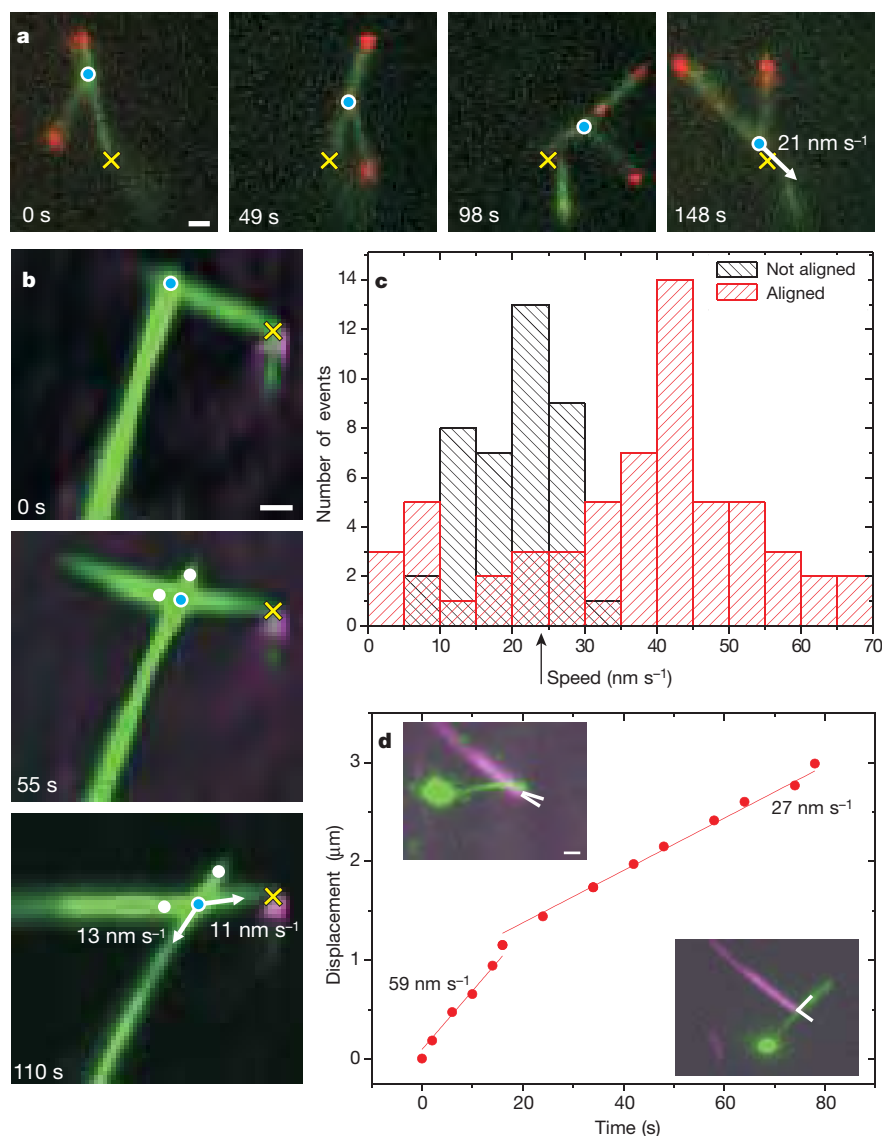


Figure 2 Eg5 moves on both filaments. **a**, Example of motility along one filament. The plus end of one microtubule (blue dot) moved towards the plus end of the other microtubule (yellow cross) at 21 nm s^{-1} . **b**, Sliding between two crossed bundles of microtubules, the far ends of which were fixed to axonemes. The blue dot marks the actual contact point, white dots mark the initial contact point on both microtubules, demonstrating simultaneous movement along both microtubules. **c**, Histograms of all measured velocities in aligned ($n = 60$) and not aligned ($n = 40$) geometries. For each microtubule,

movement in either of the two configurations was counted as a separate event (microtubule on axoneme: $n = 67$, microtubule on microtubule: $n = 33$). The arrow marks average speed in surface gliding assays. **d**, Transition from aligned (top left inset) to orthogonal (bottom right inset) sliding for a single microtubule (microtubule plus-end suspended by optical tweezers). White lines show the angle between interacting filaments. The displacement along the microtubule axis is plotted against time. Scale bars, $1 \mu\text{m}$. See also Supplementary Videos S4, S5 and S6.

ities for parallel or anti-parallel crosslinking from the frequencies observed, because microtubules should eventually all become trapped in the parallel state. Nevertheless, these data provide clear evidence that bipolar Eg5 can crosslink microtubules in all orientations, and that it moves with orientation-independent speed relative to each microtubule.

The majority of moving microtubules continued sliding when their leading minus-end reached the end of the axonemes. To our surprise, many did not release from the axonemes once the trailing plus-end reached the plus-end of the axonemes. Microtubules usually stayed attached beyond our typical recording times of minutes, while thermally pivoting around the axoneme plus-end (Fig. 3). While scanning a sample, many microtubules were found to be thermally pivoting, attached with only one end to an axoneme end. In most cases only a single microtubule was attached to an axoneme; occasionally two or three were bound. This was probably

limited by the low microtubule concentration used rather than by the binding capacity of the ends. End attachment was maintained even at salt concentrations as high as 150 mM potassium or sodium acetate.

The end-tethering of microtubules suggests that motors remain attached to microtubule ends for considerably longer than their typical ATP-turnover time. At the motor concentrations used in our experiments, it is likely that the end-tethering was caused by more than one motor. The large rotational freedom in the connection, on the other hand, implies that the crossbridges were formed by at most very few motors¹⁶ (Fig. 3, inset). The observed bundling and alignment of microtubules is evidence for binding of motors all along the filaments and therefore excludes a specific affinity of Eg5 for microtubule ends only, which has been observed for members of the kinesin-13 family¹⁷. The simplest explanation for our data is that there is an additional, ATP-independent binding mode that is long-

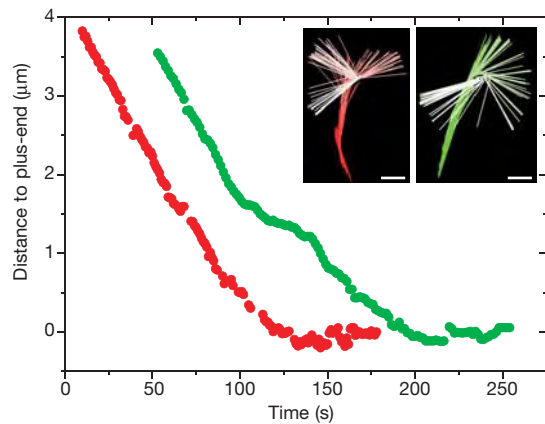


Figure 3 Eg5 can keep microtubule ends crosslinked. Two representative plots of the distance between microtubule and axoneme plus-ends against time. Sliding stops at axoneme ends but microtubules stay attached. Inserts show the corresponding superimposed video frames of the sliding microtubules, showing pivoting around the axoneme end (progression of time is colour-coded from dark to light). Scale bars, 1 μm . Time-step between frames: red, 1 s; green, 1.5 s. See also Supplementary Video S1.

range enough to allow free sliding along the track, but prevents release at the end. This is consistent with evidence for a low-friction binding mode of dimeric Eg5 in the presence of the drug monastrol¹⁸. An electrostatic microtubule-binding motif has been identified in the related BimC motor¹⁹, but is absent in Eg5. The ability of Eg5 to tether opposing microtubule plus-ends might have a role in the mitotic spindle, particularly for overlapping interpolar microtubules, and contribute to the dynamics and targeting of the motor¹⁴.

We have shown that bipolar Eg5 can (probably in a non-clustered form) slide anti-parallel microtubules apart. It is difficult to reconcile this finding with tetrameric Eg5 being non-processive, as has been reported for a truncated (dimeric) construct²⁰ and suggested from indirect experiments in spindles assembled in cell-free extracts¹². Transmission of force from one microtubule to the other strictly requires a motor to be simultaneously attached via both ends if the tetramer is not part of a physically linked cluster (in contrast with muscle myosin, for example). Non-processive motors undergo only one catalytic cycle per diffusional encounter with the filament, and are usually bound to the filament for a small fraction of their total cycle time (the duty ratio). For such a motor, the probability of simultaneous binding to two microtubules might be extremely low, because it scales with the square of the duty ratio. The majority of motors would at any point in time be bound to only one track and consume ATP without transmitting force. In contrast, a processive motor would remain attached for many cycles. Therefore, the chance that it crosslinks two microtubules is large and efficiency would be high. Thus, it seems likely that full-length Eg5 is a doubly processive kinesin (with both ends of the tetramer being processive on separate microtubules).

Other mitotic kinesins can also slide microtubules apart, although the exact mechanisms differ or are unknown. Minus-end-directed dimeric ncd (kinesin-14) is thought to stably anchor onto one filament with an ATP-independent binding site in the tail and generate force non-processively along the other microtubule²¹. The speed doubling we found with Eg5 clearly excludes an ncd-like anchored mechanism. The tetrameric kinesin-6 family member MKLP1 (ref. 22) slides anti-parallel microtubules apart²³, but it is not known whether motility is generated along both tracks simultaneously, as we have shown here for Eg5.

Our findings suggest specific roles for Eg5 in spindle morphogenesis. First, we provide direct evidence that Eg5 can push anti-

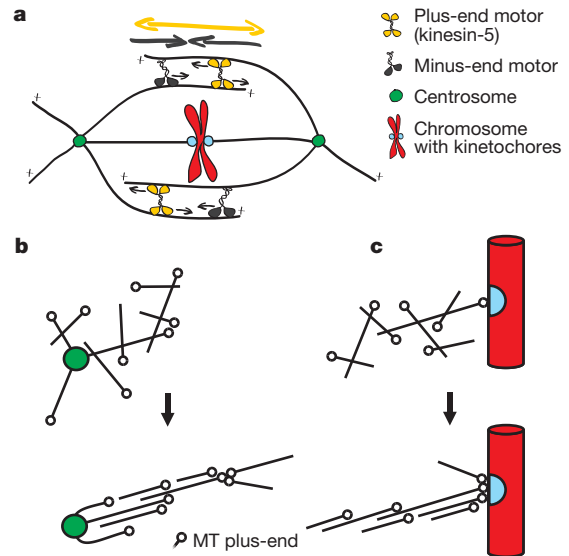


Figure 4 Model for the contribution of Eg5 to mitotic spindle morphogenesis. **a**, Push–pull model, with opposing motors acting on microtubules. The plus-end motors can be kinesin-5, and the minus-end motors dynein or kinesin-14. **b**, **c**, Potential roles of Eg5 near centrosomes (**b**) and chromosomes (**c**). Eg5 could recruit and sort microtubules (MT), resulting in bundles of parallel microtubules anchored either to chromosomes or centrosomes.

parallel microtubules apart. This is a long-standing but as yet unproven hypothesis and is a primary function assigned to bipolar motors in the push–pull muscle-like model for the spindle (Fig. 4a). Furthermore, in a cloud of short microtubules, Eg5 can first condense the microtubules into aligned bundles and subsequently sort them apart according to orientation, possibly aligning the plus-ends of parallel microtubules. Such a process might operate around the centrosomes in the initial phase of spindle morphogenesis (Fig. 4b), and might also contribute to the formation of microtubule bundles important for chromosome–spindle attachment (Fig. 4c).

We anticipate our assay to be a starting point for more sophisticated *in vitro* models of mitotic spindles. For example, the individual and combined action of multiple mitotic motors could be tested, including minus-end-directed motors opposing Eg5 motility. Furthermore, Eg5 inhibition is a major target of anti-cancer drug development, and a well-defined and quantitative assay for motor function will be relevant for such developments. □

Methods

Protein purification

Full-length *Xenopus* Eg5 with an amino-terminal 6-histidine tag was expressed in insect cells, purified as described¹⁴ and stored at -80°C . For controls, tetrameric Eg5 was further purified using gel filtration on a Superose 6 column¹⁴ and used immediately (without freezing). Truncated Eg5 with amino acids 1–591 and a carboxy-terminal polyhistidine tag was expressed in bacteria and purified as described¹⁵. Axonemes were purified from sea urchin sperm essentially as described²⁴. Fluorescent axonemes were prepared by incubating with mono-reactive NHS-Ester Cy5 (Amersham Biosciences), followed by three rounds of centrifugation and resuspension.

Tubulin was purified from porcine brain by two cycles of assembly and disassembly, followed by chromatography on phosphocellulose²⁵. Microtubules were polymerized by incubating 7.5 μM unlabelled tubulin and 2.5 μM rhodamine-labelled tubulin (Cytoskeleton) in the presence of 1 mM GMPCPP (Guanosine-5'[(α , β)-methylene]triphosphate, Jena Bioscience) and 2 mM dithiothreitol at 35°C for 15 min. For polarity-marked microtubules, non-fluorescent seeds were polymerized from 13.5 μM unlabelled tubulin. These were then labelled with Cy5 and stored at -80°C . Cy5-labelled seeds were elongated by adding 0.3 μM rhodamine-labelled tubulin, 4.5 μM unlabelled

tubulin, 4.5 μM N-ethylmaleimide (NEM)-tubulin²⁶, 1 mM GTP and 2 mM dithiothreitol, followed by incubation at 35 °C for 20 min. NEM-tubulin was added to inhibit minus-end growth, resulting in rhodamine-microtubule growth from the plus-end of the Cy5-labelled seeds only. Polymerized microtubules were stabilized using 10 μM paclitaxel (Sigma).

In vitro assays

Assays were performed at 21 °C using an epi-illuminated wide-field fluorescence microscope capable of optical trapping and laser-darkfield detection as described elsewhere²⁷. Excitation light (532 nm) was coupled into the objective with a polychromatic dichroic mirror (532/633PC, Chroma), allowing combined excitation with a 633-nm HeNe laser (Coherent) as well as transmission darkfield imaging with a 650-nm diode laser (Roithner Lasertechnik). Emission was first short-pass filtered (Chroma E750sp), further filtered with holographic notch filters against 532-nm and 633-nm light (Kaiser Optical Systems HNPF-532 and HNPF-632.8), separated with a dichroic mirror (645DCXRlp, Chroma) and imaged side-by-side on the CCD-camera, allowing simultaneous imaging of rhodamine emission and either darkfield light or Cy5 emission. For darkfield detection of axonemes, we used a darkfield condenser (Nikon) in combination with a decreased numerical aperture (NA) of the objective to block illumination light. Trapping experiments required full NA and hence we used fluorescence imaging of axonemes in those experiments. The trapping laser was used at 850-nm at a power of typically 150 mW in the sample.

Hydrophobic sample chambers were assembled by joining dimethyl-dichlorosilane-treated slides and coverslips using two layers of double-stick tape (~150- μm inner height). A chamber was first incubated with axonemes in PEM80 (80 mM PIPES, 1 mM EGTA, 2 mM MgCl_2 , pH 6.8) for 10 min, then blocked by washing with 5 volumes of PEM80 with amphiphilic copolymers (0.2% (w/v) Pluronic F108, BASF) and 0.1% (w/v) polyethylene-block-poly(ethylene glycol) (PEPEG, Aldrich). Finally, the chamber was filled with the motility sample, consisting of PEM80 with microtubules, 6–12 $\mu\text{g ml}^{-1}$ Eg5, 6–10 mM ATP, 4 mM dithiothreitol, 25 mM glucose, 20 $\mu\text{g ml}^{-1}$ glucose oxidase, 35 $\mu\text{g ml}^{-1}$ catalase, 10 μM paclitaxel, 0.2% (w/v) Pluronic F108, and in some cases 0.1% (w/v) PEPEG and/or 0.1% (w/v) methyl cellulose (Fluka). For trapping experiments, 1- μm fluorescent silica beads (Kisker Biotech) treated with anti- α -tubulin antibodies (Sigma) were added, and methyl cellulose was left out. Measurements at high salt were performed with 150 mM sodium acetate in PEM80 or 150 mM potassium acetate in PEM12 (equivalent to PEM80, except with 12 mM PIPES), together with a two- to threefold increase in Eg5 concentration. This was necessary to compensate for a lower probability of microtubule capture at higher salt. For controls without axonemes, sample mix was used with an increased concentration of both microtubules and Eg5 (full-length or truncated) to enhance crosslinking in solution.

Assays with polarity-marked microtubules were performed in sample chambers of 5- μm inner height (set by the addition of spacer beads, Seradyn) to confine microtubules near the focal plane. Both surfaces were pre-incubated with the copolymers and dried, after which the sample mix was put onto one surface and covered with the other. Surface blocking was not perfect in these samples, but good enough to keep sufficient microtubules and motors in solution. Imperfect blocking was an advantage because, for stable imaging, we could choose events in which at least one microtubule was partially stuck to the surface.

Surface gliding assays were performed under the same buffer conditions but without surface blocking.

Image analysis

Digital images were analysed using custom-written routines in LabVIEW (National Instruments). Speeds were determined by measuring displacement relative to a reference point on one of the filaments. For Fig. 3, every 5–10 frames were averaged to reduce the number of frames. Microtubules were automatically tracked and the position of the trailing end was measured relative to its final position.

Received 13 January; accepted 2 March 2005; doi:10.1038/nature03503.

- Scholey, J. M., Brust-Mascher, I. & Mogilner, A. Cell division. *Nature* **422**, 746–752 (2003).
- Compton, D. A. Spindle assembly in animal cells. *Annu. Rev. Biochem.* **69**, 95–114 (2000).
- Kashina, A. S., Rogers, G. C. & Scholey, J. M. The bimC family of kinesins: essential bipolar mitotic motors driving centrosome separation. *Biochim. Biophys. Acta* **1357**, 257–271 (1997).
- McIntosh, J. R., Hepler, P. K. & Vanwie, D. G. Model for mitosis. *Nature* **224**, 659–663 (1969).
- Sharp, D. J., Rogers, G. C. & Scholey, J. M. Microtubule motors in mitosis. *Nature* **407**, 41–47 (2000).
- Brust-Mascher, I. & Scholey, J. M. Microtubule flux and sliding in mitotic spindles of *Drosophila* embryos. *Mol. Biol. Cell* **13**, 3967–3975 (2002).
- Kashina, A. S. *et al.* A bipolar kinesin. *Nature* **379**, 270–272 (1996).
- Sharp, D. J. *et al.* The bipolar kinesin, KLP61F, cross-links microtubules within interpolar microtubule bundles of *Drosophila* embryonic mitotic spindles. *J. Cell Biol.* **144**, 125–138 (1999).
- Sawin, K. E., Leguellec, K., Philippe, M. & Mitchison, T. J. Mitotic spindle organization by a plus-end-directed microtubule motor. *Nature* **359**, 540–543 (1992).
- Wilde, A. *et al.* Ran stimulates spindle assembly by altering microtubule dynamics and the balance of motor activities. *Nature Cell Biol.* **3**, 221–227 (2001).
- Mayer, T. U. *et al.* Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotypic-based screen. *Science* **286**, 971–974 (1999).
- Miyamoto, D. T., Perlman, Z. E., Burbank, K. S., Groen, A. C. & Mitchison, T. J. The kinesin Eg5 drives poleward microtubule flux in *Xenopus laevis* egg extract spindles. *J. Cell Biol.* **167**, 813–818 (2004).
- Kwok, B. H., Yang, J. G. & Kapoor, T. M. The rate of bipolar spindle assembly depends on the microtubule-gliding velocity of the mitotic kinesin Eg5. *Curr. Biol.* **14**, 1783–1788 (2004).
- Kapoor, T. M. & Mitchison, T. J. Eg5 is static in bipolar spindles relative to tubulin: evidence for a static spindle matrix. *J. Cell Biol.* **154**, 1125–1133 (2001).
- Kapoor, T. M. & Mitchison, T. J. Allele-specific activators and inhibitors for kinesin. *Proc. Natl Acad. Sci. USA* **96**, 9106–9111 (1999).

- Howard, J., Hudspeth, A. J. & Vale, R. D. Movement of microtubules by single kinesin molecules. *Nature* **342**, 154–158 (1989).
- Desai, A., Verma, S., Mitchison, T. J. & Walczak, C. E. Kin I kinesins are microtubule-destabilizing enzymes. *Cell* **96**, 69–78 (1999).
- Crevel, I. M., Alonso, M. C. & Cross, R. A. Monastrol stabilises an attached low-friction mode of Eg5. *Curr. Biol.* **14**, R411–R412 (2004).
- Stock, M. F., Chu, J. & Hackney, D. D. The kinesin family member BimC contains a second microtubule binding region attached to the N terminus of the motor domain. *J. Biol. Chem.* **278**, 52315–52322 (2003).
- Crevel, I., Lockhart, A. & Cross, R. A. Kinetic evidence for low chemical processivity in ncd and Eg5. *J. Mol. Biol.* **273**, 160–170 (1997).
- McDonald, H. B., Stewart, R. J. & Goldstein, L. S. B. The kinesin-like ncd protein of *Drosophila* is a minus end-directed microtubule motor. *Cell* **63**, 1159–1165 (1990).
- Chui, K. K. *et al.* Roles of two homotetrameric kinesins in sea urchin embryonic cell division. *J. Biol. Chem.* **275**, 38005–38011 (2000).
- Nislow, C., Lombillo, V. A., Kuriyama, R. & McIntosh, J. R. A plus-end-directed motor enzyme that moves antiparallel microtubules *in vitro* localizes to the interzone of mitotic spindles. *Nature* **359**, 543–547 (1992).
- Gibbons, I. R. & Fronk, E. Latent adenosine-triphosphatase form of dynein-1 from sea urchin sperm flagella. *J. Biol. Chem.* **254**, 187–196 (1979).
- Williams, R. C. & Lee, J. C. Preparation of tubulin from brain. *Methods Enzymol.* **85**, 376–385 (1982).
- Hyman, A. A. Preparation of marked microtubules for the assay of the polarity of microtubule-based motors by fluorescence. *J. Cell Sci.* **14** (suppl.), 125–127 (1991).
- van Dijk, M. A., Kapitein, L. C., van Mameren, J., Schmidt, C. F. & Peterman, E. J. G. Combining optical trapping and single-molecule fluorescence spectroscopy: Enhanced photobleaching of fluorophores. *J. Phys. Chem. B* **108**, 6479–6484 (2004).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank I. Schaap for purifying tubulin, S. Calmat and J. Hendriks for assistance with FPLC chromatography, J. van Mameren for help with software, and M. Korneev, M. Mazur and K. Zaborcka for help with surface chemistry and motility controls. L.C.K. and E.J.G.P. are supported by a VIDI fellowship to E.J.G.P. from the Research Council for Earth and Life Sciences (ALW), with financial aid from the Netherlands Organization for Scientific Research (NWO). T.M.K., B.H.K., J.H.K. are grateful to the NIH/NIGMS for support. Additional support was provided by the Foundation for Fundamental Research on Matter (C.E.S.) and a Research Grant from the Human Frontier Science Program (C.E.S. and T.M.K.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.J.G.P. (erwinp@nat.vu.nl) or T.M.K. (kapoor@mail.rockefeller.edu).

A synthetic gene–metabolic oscillator

Eileen Fung^{1,2}, Wilson W. Wong¹, Jason K. Suen¹, Thomas Bulter¹, Sun-gu Lee¹ & James C. Liao^{1,2}

¹Department of Chemical Engineering and ²Biomedical Engineering Interdepartmental Program, University of California–Los Angeles, Los Angeles, California 90095, USA

Autonomous oscillations found in gene expression and metabolic, cardiac and neuronal systems^{1–4} have attracted significant attention both because of their obvious biological roles and their intriguing dynamics. In addition, *de novo* designed^{5–12} oscillators^{13,14} have been demonstrated, using components that are not part of the natural oscillators. Such oscillators are useful in testing the design principles and in exploring potential applications not limited by natural cellular behaviour¹⁵. To achieve transcriptional and metabolic integration characteristic of natural oscillators, here we designed and constructed a synthetic circuit in *Escherichia coli* K12, using glycolytic flux to generate oscillation through the signalling metabolite acetyl phosphate. If two metabolite pools are interconverted by two enzymes that are placed under the transcriptional control of acetyl phosphate, the system oscillates when the glycolytic rate exceeds a critical value. We used bifurcation analysis to identify the boundaries of

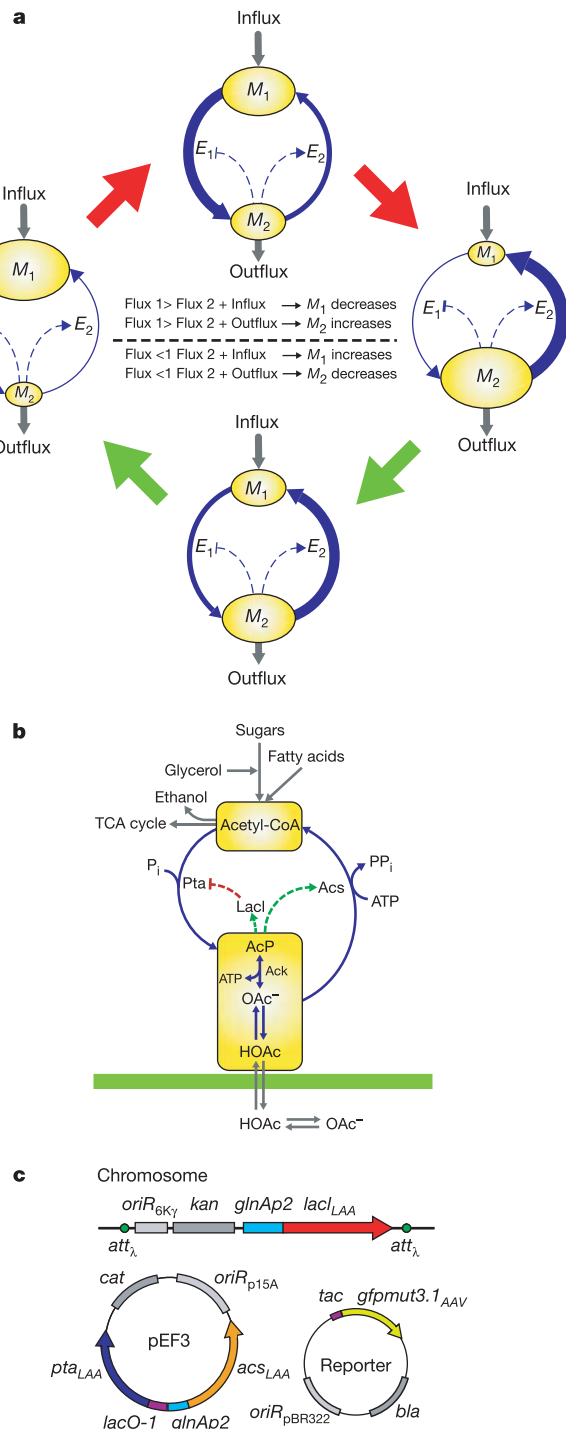


Figure 1 The design and construction of the metabolator. **a**, Conceptual diagram of the oscillatory dynamics, highlighting the two metabolite pools (M_1 and M_2) and their controls. Solid lines indicate metabolic fluxes. Dashed lines indicate positive (arrow) and negative (blunt bar) transcriptional or translational regulation. Input metabolic flux increases M_1 . E_1 drives M_1 to M_2 . Accumulation of M_2 represses E_1 and upregulates E_2 expression. Low E_1 and high E_2 levels drive M_2 to M_1 and a new cycle begins. **b**, Biological realization of the conceptual design in **a**. The yellow boxes highlight the two metabolic pools, M_1 and M_2 . Ack, acetate kinase; AcP, acetyl phosphate; Acs, acetyl-CoA synthetase; OAc⁻, acetate; Pta, phosphate acetyltransferase. **c**, *E. coli* strain and plasmid constructs. The *lacI* gene under the control of the *glnAp2* promoter was inserted into the *E. coli* chromosome at the lambda attachment site (*att_λ*) (see Methods). A degradation tag was fused to the 3'-end of the gene to reduce the stability of LacI (*lacI_{LAA}*). The pEF3 plasmid expresses low-stability Pta (*pta_{LAA}*) and Acs (*acs_{LAA}*), which are controlled by the *lacO-1* and *glnAp2* promoters, respectively. The reporter plasmid expresses an intermediate-stability GFP variant (*gfpmut3.1_{AAV}*) under the control of the *tac* promoter. *kan*, *cat* and *bla* are genes for resistance to kanamycin, chloramphenicol and ampicillin, respectively.

oscillation, and verified these experimentally. This work demonstrates the possibility of using metabolic flux as a control factor in system-wide oscillation, as well as the predictability of a *de novo* gene-metabolic circuit designed using nonlinear dynamic analysis.

Our design of this oscillatory circuit, termed the metabolator, integrates transcriptional regulation with metabolism. This is in contrast with the yeast glycolytic oscillator^{1,4}, which does not involve transcriptional regulation, as well as previous synthetic gene expression oscillators, which were independent of metabolism^{13,14}. The conceptual design of the metabolator consists of a flux-carrying network with two interconverting metabolite pools (M_1 and M_2) catalysed by two enzymes (E_1 and E_2), the expressions of which are negatively and positively regulated by M_2 , respectively (Fig. 1a). In the first stage, when the level of M_2 is low, E_1 is expressed and E_2 is not. A high input metabolic flux rapidly drives M_1 to M_2 . Accumulation of M_2 represses E_1 and upregulates E_2 . When the backward reaction rate exceeds the sum of the forward reaction rate and the output rate, the level of M_2 decreases and the

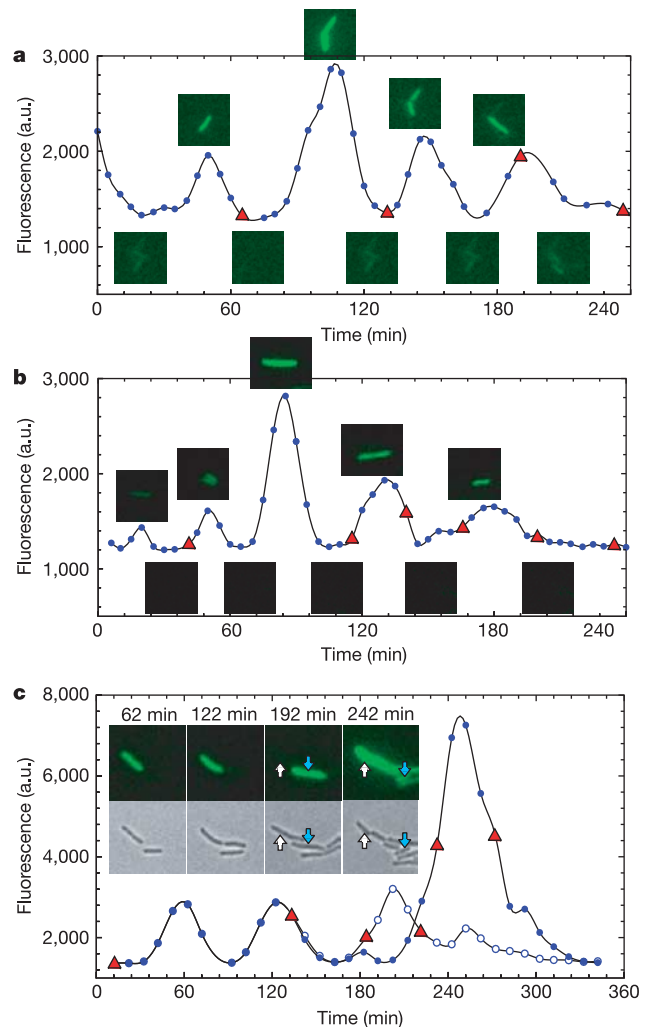


Figure 2 Oscillation dynamics of the metabolator in glucose. In each panel, red triangles denote cell division. Fluorescence intensity shown in arbitrary units. Inserts show the fluorescence intensities at peaks and valleys during oscillation. **a**, **b**, Representative single cell dynamics evaluated at 30 °C from ten independent experiments. **c**, Representative oscillation dynamics of two sibling cells evaluated at 37 °C from four independent experiments. The arrows (blue or white) in inserts point to sibling cells after cell division. Comparison of sibling cells shows skipped or delayed peaks (compare filled and open circles), which might be partially explained by uneven distribution of cellular materials.

M_1 level increases. E_1 is then expressed again and E_2 is degraded, returning the circuit to the first stage. On the other hand, if the input flux is low, M_2 does not accumulate sufficiently fast to cause a large swing in gene expression, and a stable steady state can be reached. This design allows metabolic physiology to influence gene expression cycles, a characteristic seen in circadian regulation.

Such a conceptual design was realized using the acetate pathway in *E. coli* (Fig. 1b). The M_1 pool is acetyl coenzyme A (acetyl-CoA) and the M_2 pool consists of acetyl phosphate (AcP), acetate (OAc^-) and the protonated form of acetate (HOAc). Acetyl-CoA is a metabolic product of sugar, fatty acids and some amino acids, and is the entry point into the tricarboxylic acid (TCA) cycle. Acetyl-CoA is converted to acetyl phosphate in *E. coli* by phosphate acetyltransferase (Pta), which corresponds to enzyme E_1 in Fig. 1a, and then to acetate by acetate kinase (Ack). The protonated form of acetate is permeable across the cell membrane. Under aerobic conditions, acetyl-CoA is further oxidized in the TCA cycle. The remaining flux goes to produce either acetate or ethanol. In wild-type *E. coli*, the enzyme acetyl-CoA synthetase (Acs) is induced in the presence of acetate¹⁶. However, such induction is under catabolite repression by glucose in the wild-type strain so as to avoid futile cycling. In our design, acetyl-CoA synthetase is used as enzyme E_2 in Fig. 1a. Thus, both phosphate acetyltransferase and acetyl-CoA synthetase need to be re-wired to respond to the M_2 pool.

Acetyl phosphate in the M_2 pool is used as a signalling molecule that activates the *glnAp2* promoter through phosphorylation of the nitrogen regulator I (NR_I or NtrC) in strains lacking the sensor, nitrogen regulator II (NR_{II} or NtrB), which is encoded by the *glnL* gene^{5,11,17}. We used a ΔglnL strain as the host and placed the *acs* gene under the control of *glnAp2* promoter in a plasmid. When acetyl phosphate builds up, it activates the *glnAp2* promoter, which controls the expression of *acs*. At the same time, acetyl phosphate represses *pta* expression indirectly: a *glnAp2* promoter is used to control the production of the *lac* repressor (LacI), which in turn represses the expression of the *pta* gene from the *lacO-1* promoter¹⁸. The three proteins Acs, LacI and Pta are each fused with a

degradation tag at their carboxy terminals so as to reduce their half-life¹⁹.

We constructed the metabolator strain and plasmid according to the design shown in Fig. 1c. The host strain used was BW18793 (ref. 20), which contains *glnL* and *pta* null mutations and *lac* deletion. The *lacI* gene, controlled by the *glnAp2* promoter, was integrated into the chromosome using the CRIM method²¹. The *pta* and *acs* genes were cloned into plasmid pACYC184, under the control of *lacO1* and *glnAp2* promoters, respectively. To monitor the behaviour of the metabolator, a green fluorescent protein (GFP) fused with a degradation tag¹⁹ was ligated downstream of a plasmid-borne *LacI*-repressible *tac* promoter, to be used as a readout. We demonstrated (see Supplementary Information) that the oscillation behaviour was insensitive to the degradation rate of GFP, which affects the phase but not the period of oscillation. Therefore, the period observed from the GFP readout reflects the dynamics of the system.

We monitored the GFP expression of individual cells grown on a thin agar gel with glucose and casamino acids under a fluorescence microscope at 30 °C (Fig. 2a, b) and 37 °C (Fig. 2c). About 60% of the initial colonies oscillate with varying amplitude and a period of 45 ± 10 min (mean $\pm$ s.d. for a total of 85 cells in 10 experiments), for a total duration of at least 4 h. The cell divisions (generation time ~ 60 min) appeared to be uncorrelated with the oscillation cycle, suggesting that the metabolator is responsible for the oscillation and is independent of the cell cycle. The baseline for the fluorescence and frequency of oscillation remained reasonably constant (Fig. 2a–c) but the amplitude varied. After division, the daughter cells commonly show uneven fluorescence (Fig. 2c), possibly because of stochastic variation in distributing cellular materials. Acetate was secreted from the cell, but was later shown to suppress rather than synchronize oscillation in both liquid and agar cultures.

To analyse the system quantitatively, we constructed a nonlinear ordinary differential equation (ODE) model on the basis of a dynamic balance of biomolecules. Enzyme kinetics were described by Michaelis–Menten kinetics and gene expression kinetics were

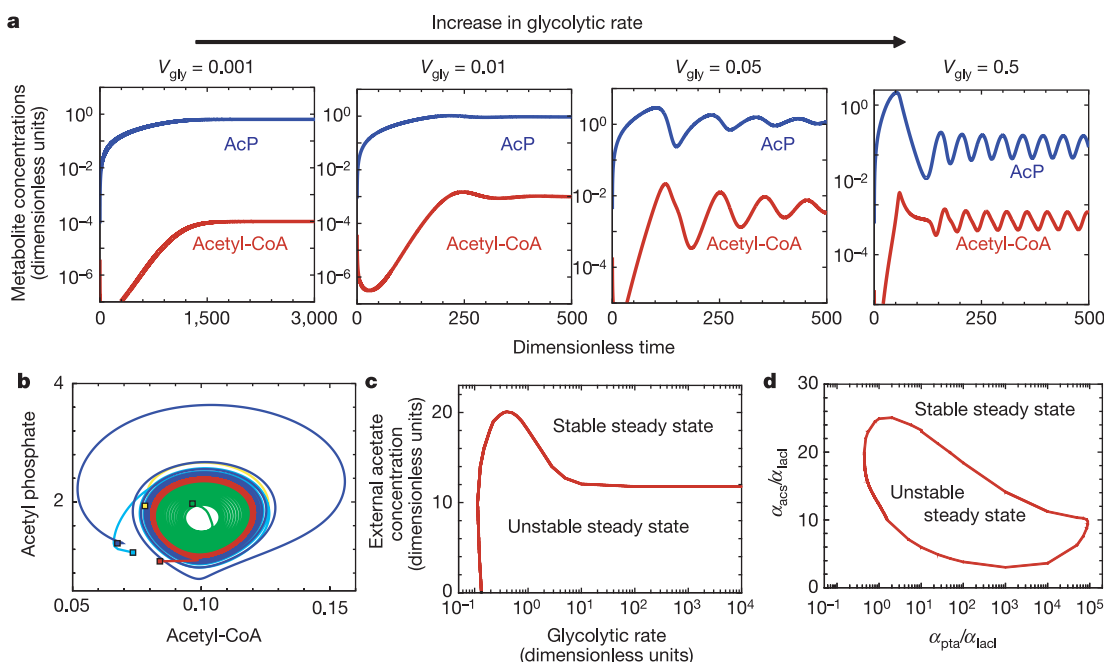


Figure 3 Computational characterization of the metabolator. **a**, The metabolator is prone to oscillate at increasing glycolytic rates (V_{gly}). V_{gly} in the four panels (from left to right) are 0.001, 0.01, 0.05 and 0.5. **b**, Phase plots constructed by perturbing steady state solution at $V_{\text{gly}} = 1$ reveal the oscillatory dynamics is limit cycle oscillation. Squares represent the initial state of the system. **c**, Phase diagram of glycolytic rate and external acetate

concentration reveals the flux-sensitive oscillation behaviour of the metabolator. **d**, A phase diagram suggests that specific combinations of three proteins are necessary to support oscillation with $V_{\text{gly}} = 10$. α_i is related to the gene copy number for the synthesis of protein i (where i is LacI, Pta or Acs).

described using the Hill equation (see Supplementary Information). Numerical simulations show that with increased glycolytic flux, the system is prone to oscillate (Fig. 3a) and approach a limit cycle (Fig. 3b) regardless of the initial condition. Floquet analysis shows that the limit cycle is stable (see Supplementary Information). In addition, Hopf bifurcation analysis shows that the system operates at a stable, steady state at low glycolytic fluxes, but becomes oscillatory when the glycolytic flux increases above a critical point (Fig. 3c). On the other hand, increasing the external acetate concentration stabilizes the system and suppresses oscillation.

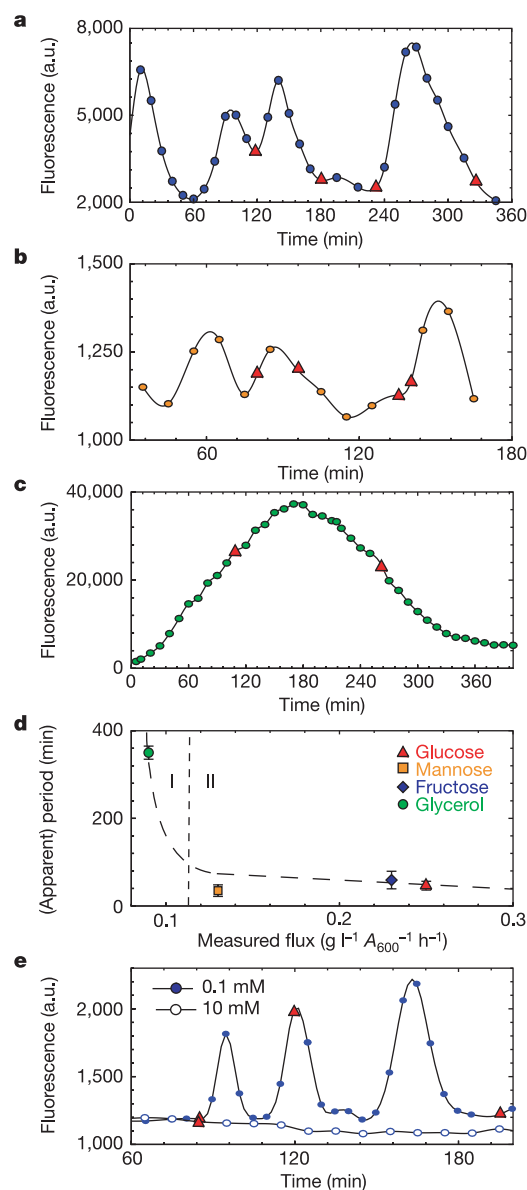


Figure 4 Flux-sensitive oscillation. Red triangles in **a–c** and **e** denote cell division. **a–c**, Dynamics using fructose (**a**), mannose (**b**) or glycerol (**c**) as a carbon source. **d**, Effect of glycolytic flux on period (fructose, glucose and mannose) and apparent period (glycerol). Apparent period is the time between the first and second valleys in a damped oscillation. The (apparent) periods are taken as the average of 85 (glucose), 32 (fructose), 18 (mannose) or 5 (glycerol) cells from three independent biological repeats. Glycolytic flux is evaluated by averaging at least three independent measurements. Error bars show standard deviation between repeats. Dashed line shows the theoretical correlation between (apparent) period and glycolytic flux. Region I denotes stable steady state and region II denotes a periodic state. **e**, Effect of acetate on metabolator dynamics. 10 mM external acetate suppressed oscillation but 0.1 mM did not.

To verify our theoretical predictions, we changed the glycolytic rate by varying the carbon source from glucose to less favourable carbon sources (fructose, mannose or glycerol) under otherwise identical conditions. Glycolytic flux decreased in the following order (from highest flux to lowest flux): glucose, fructose, mannose and glycerol (Fig. 4d). High glycolytic rates from glucose, fructose, and mannose supported oscillation (Figs 2 and 4a, b) but a low glycolytic rate from glycerol did not (Fig. 4c). Furthermore, addition of 10 mM acetate in the medium abolished oscillation (Fig. 4e), whereas 0.1 mM acetate still supported periodic behaviour. These data corroborate the results shown in Fig. 3c and demonstrate the role of glycolytic flux and acetate concentration on the dynamic behaviour of the metabolator.

We also used the model to explore the design parameter space of the metabolator. In particular, bifurcation analysis using the gene copy numbers of *acs* and *pta* relative to the chromosomal copy of *lacI* as parameters shows that the oscillation is characterized by an island defined by specific combinations of relative Acs and Pta levels (Fig. 3d). This analysis suggests that the relative gene copy number is an important determinant of oscillatory behaviour. This prediction is consistent with experimental results: when Acs was omitted in the metabolator, the system ceased to oscillate. Furthermore, the island of oscillation might partially explain the relatively few cycles of oscillation observed, as cell growth might move the system out of the oscillation region when the intracellular and extracellular environment changes. With biological stochasticity, some cells might operate outside of the oscillation island, and thus not all cells show oscillation.

The metabolator illustrates the use of metabolic fluxes to control biological oscillation. Specifically, oscillation is generated by the glycolytic flux leading to acetyl-CoA. Once the glycolytic flux exceeds a critical value, the system ‘chokes’ and begins to oscillate. However, this oscillation can be suppressed by imposing a sufficiently high concentration of acetate, which maintains an acetyl phosphate level outside of the dynamic range of the *glnAp2* promoter response. In the agar culture, accumulation of acetate produced from the cell eventually moves the system out of the oscillation region. Long-term observation of oscillation thus requires perfusion of the system.

The intrinsic noise in transcriptional regulation and the discrete nature of biological processes might explain some of the variations observed experimentally. For example, the amplitude of oscillation varied considerably. This variation was uncorrelated with cell division, and therefore, the intrinsic noise in gene expression processes appears to be responsible for determining oscillatory amplitude. Simulation results accounting for stochastic variations in the expression and degradation kinetics appear to explain the observed data (Supplementary Fig. S1). Upon cell division, sibling cells showed phase-shifts or peak omissions, possibly due to uneven distribution of cellular material that sets the daughter cells to a different phase of oscillation. In some cases, some daughter cells cease to oscillate, owing to a dramatic change in the relative levels of LacI, Acs and Pta.

The unique design of the metabolator represents a fusion of metabolic and transcriptional oscillation, and allows external carbon sources to influence oscillation. Natural metabolic oscillators, such as yeast glycolysis¹, are mainly based on allosteric enzyme regulation, which has a timescale of about 5 min. Purely transcriptional oscillators typically show a period of hours, with previous examples including 2.5 h (ref. 13) and 10 h (ref. 14). As a hybrid of metabolic and transcriptional oscillators, the metabolator shows an intermediate period (45 min); this is consistent with our expectations. The metabolator involves both negative feedback (E_1) and positive feed-forward (E_2) regulation. Negative feedback is typically used to stabilize a system, however, it tends to be unstable if the time lag in feedback increases. Negative feedback systems in a loop design can also generate oscillation¹³.

The function of metabolator is somewhat analogous to the circadian clock, in that both are integrated with metabolic physiology and respond to external stimuli (carbon sources in the former, light in the latter). The close involvement of metabolism in circadian regulation has been documented. For example, metabolic entrainment^{22,23} is recognized as serving important physiological roles. Haem biosynthesis is known to be regulated by the circadian clock²⁴ and acetyl-CoA has been shown to participate in the conversion of serotonin to *N*-acetylserotonin, which is responsible for the melatonin circadian rhythm in vertebrates²⁵. The metabolator might provide useful insights into the operating principles of physiologically integrated natural oscillators. Moreover, the role of metabolic flux in controlling oscillation demonstrated here suggests that network connectivity alone is insufficient to determine cellular behaviour. In addition to mapping the regulatory wiring diagram in cells, dynamic characterization will also be essential. □

Methods

Bacterial strains and plasmids

The *lacI* gene under the control of the *glnAp2* promoter was inserted into the chromosome of *E. coli* BW18793 (ref. 20) at the lambda attachment site (*att_λ*) using the CRIM method²¹ to produce the host strain. The *glnAp2*, *tac* and *lacO-1* promoters were obtained from p21DI (ref. 5), pKK223-3 (ref. 26) and pZE12 (ref. 18), respectively. The *pta* gene was cloned by polymerase chain reaction from the *Bacillus subtilis* chromosome. The *acs* and *lacI* genes were cloned from the *E. coli* MG1655 chromosome. For both *pta* and *lacI*, the GTG start codon was changed to ATG using their respective 5' primers. The 11-amino-acid *ssrA* tag²⁷ was added to the genes of the metabolator via the 3' primers. The resulting protein, carrying the peptide tag, is recognized and targeted for degradation by tail-specific Tsp proteases. The pEF3 plasmid was constructed in two sequential cloning steps from intermediate plasmids containing the two genes into pACYC184. The reporter plasmid was constructed by cloning *gfpmut3.1* with the degradation tag into pKK223-3.

Microscopy and analysis of gene expression

Cells were grown at 37 °C in M9 media with 0.4% of an appropriate carbon source (glucose, fructose, mannose or glycerol), 0.1% casamino acids and antibiotics (5 µg ml⁻¹ kanamycin, 25 µg ml⁻¹ chloramphenicol and 100 µg ml⁻¹ ampicillin) and harvested when the absorbance at 600 nm reached 0.5–1.3. Cells were concentrated by centrifugation and the resulting pellet was resuspended in 1 ml growth medium for preparation of microscopy slides. Time-lapse microscopy was performed using a Nikon TE2000-S microscope with a × 60 DIC oil immersion objective. Images were captured using a 512 × 512-pixel cooled CCD camera controlled through Metamorph software. An agar pad was created as described in ref. 28. The temperature of the samples was maintained at approximately 30 °C by an objective heater. Brightfield (0.1 s) and epifluorescence (0.1 s) images were captured every 5 min, with both light sources shuttered between exposures. All image analysis was done using Metamorph. Cells from the final frame of the brightfield video image were randomly chosen and manually tracked back to the first frame. In each frame of the brightfield image, regions were drawn around cells using Metamorph and transferred to the fluorescence image, and the average intensities of the regions were recorded.

Each of the fluorescence measurements for the metabolator dynamics was obtained from at least three independent cultures. For each culture, at least 60% of the cells were tracked for GFP dynamics in response to oscillation induced by glucose (*n* = 85, Fig. 2), fructose (*n* = 32, Fig. 4a), mannose (*n* = 18, Fig. 4b), glycerol (*n* = 5, Fig. 4c) or two different concentrations of acetate (*n* = 27, Fig. 4e). In addition, various negative controls were conducted. These included fluorescence measurements for the metabolator supplemented with 100 µM isopropyl-β-D-thiogalactoside (IPTG), for the host strain containing a reporter plasmid only, and for the host strain with a chromosome-integrated *pta* gene under the control of the *tac* promoter and the reporter plasmid.

Glycolytic flux measurements

The oscillation strain (without the reporter plasmid) was grown overnight in M9 media containing 0.4% of an appropriate carbon source, as described above. Cells were then collected by centrifugation and resuspended to an absorbance at 600 nm of 0.5 in fresh M9 media containing 0.1% casamino acids, antibiotics and 0.04% carbon source. Samples were collected every 30 min for supernatant and cell density measurements. Cell density was measured in triplicate using a microplate reader. Glucose and fructose concentrations were measured in triplicate with a D-glucose/D-fructose kit and glycerol was measured with a glycerol kit (R-biopharm), in accordance with the manufacturer's instructions. Mannose was measured using a dinitrosalicylic acid (DNS) assay. Carbon flux was calculated on the basis of sugar consumption.

Physicochemical modelling and numerical analysis

A nonlinear ODE-based model was developed to describe the key components of the metabolator. Exact expressions of the model are presented in the Supplementary Information. The computational approach used to elucidate the dynamics and behaviour of this model involves two components. First, direct numerical simulation was conducted to explore the nonlinear dynamics of the model (for example, Fig. 3a). The exact

parameter values used in this simulation are shown in Supplementary Table S1. Next, phase diagrams that mapped out the location of Hopf bifurcation as functions of key parameters were constructed to study the asymptotic behaviour of the model²⁹. Each phase diagram was constructed by mapping the locus of Hopf bifurcation as a function of key parameters. Specifically, the response of the steady-state solution to infinitesimal perturbations was computed by solving a linearized model about the same steady state. This linear response is determined by calculation of the eigenvalues in the linear spectrum that have the largest real part, σ . The onset of the Hopf bifurcation is determined by the moment at which the real part of a complex conjugate pair of eigenvalues of the jacobian matrix crosses zero, while the real parts of all other eigenvalues remain negative.

Received 16 November 2004; accepted 28 February 2005; doi:10.1038/nature03508.

- Goldbeter, A. *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour* (Cambridge Univ. Press, New York, 1996).
- Hess, B. Periodic patterns in biology. *Naturwissenschaften* **87**, 199–211 (2000).
- Bier, M., Teusink, B., Kholodenko, B. N. & Westerhoff, H. V. Control analysis of glycolytic oscillations. *Biophys. Chem.* **62**, 15–24 (1996).
- Bier, M., Bakker, B. M. & Westerhoff, H. V. How yeast cells synchronize their glycolytic oscillations: a perturbation analytical treatment. *Biophys. J.* **78**, 1087–1093 (2000).
- Farmer, W. R. & Liao, J. C. Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nature Biotechnol.* **18**, 533–537 (2000).
- Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
- Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
- Guet, C. C., Elowitz, M. B., Hsing, W. & Leibler, S. Combinatorial synthesis of genetic networks. *Science* **296**, 1466–1470 (2002).
- Isaacs, F. J., Hasty, J., Cantor, C. R. & Collins, J. J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl Acad. Sci. USA* **100**, 7714–7719 (2003).
- Basu, S., Mehreja, R., Thiberge, S., Chen, M. T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl Acad. Sci. USA* **101**, 6355–6360 (2004).
- Bulter, T. et al. Design of artificial cell–cell communication using gene and metabolic networks. *Proc. Natl Acad. Sci. USA* **101**, 2299–2304 (2004).
- You, L., Cox, R. S. III, Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428**, 868–871 (2004).
- Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
- Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* **113**, 597–607 (2003).
- Kobayashi, H. et al. Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl Acad. Sci. USA* **101**, 8414–8419 (2004).
- Kumari, S. et al. Regulation of acetyl coenzyme A synthetase in *Escherichia coli*. *J. Bacteriol.* **182**, 4173–4179 (2000).
- McCleary, W. R. & Stock, J. B. Acetyl phosphate and the activation of two-component response regulators. *J. Biol. Chem.* **269**, 31567–31572 (1994).
- Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).
- Andersen, J. B. et al. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol.* **64**, 2240–2246 (1998).
- Feng, J. et al. Role of phosphorylated metabolic intermediates in the regulation of glutamine synthetase synthesis in *Escherichia coli*. *J. Bacteriol.* **174**, 6061–6070 (1992).
- Haldemann, A. & Wanner, B. L. Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J. Bacteriol.* **183**, 6384–6393 (2001).
- Rutter, J., Reick, M. & McKnight, S. L. Metabolism and the control of circadian rhythms. *Annu. Rev. Biochem.* **71**, 307–331 (2002).
- Rutter, J., Reick, M., Wu, L. C. & McKnight, S. L. Regulation of clock and NPAS2 DNA binding by the redox state of NAD cofactors. *Science* **293**, 510–514 (2001).
- Kaasik, K. & Lee, C. C. Reciprocal regulation of haem biosynthesis and the circadian clock in mammals. *Nature* **430**, 467–471 (2004).
- Ferry, G. et al. Substrate specificity and inhibition studies of human serotonin N-acetyltransferase. *J. Biol. Chem.* **275**, 8794–8805 (2000).
- Brosius, J. & Holy, A. Regulation of ribosomal RNA promoters with a synthetic lac operator. *Proc. Natl Acad. Sci. USA* **81**, 6929–6933 (1984).
- Keiler, K. C., Waller, P. R. & Sauer, R. T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990–993 (1996).
- Rines, D. R., He, X. & Sorger, P. K. Quantitative microscopy of green fluorescent protein-labeled yeast. *Methods Enzymol.* **351**, 16–34 (2002).
- Iooss, G. & Joseph, D. D. *Elementary Stability and Bifurcation Theory* (Springer-Verlag, New York, 1989).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements The authors thank V. Roychowdhury and J. Bridgewater for discussions. This work was partially funded by the Center for Cell Mimetic Space Exploration, a National Aeronautics and Space Administration University Research, Engineering, and Technology Institute. J.C.L. is a member of California NanoSystems Institute and UCLA-DOE Institute for Genomics and Proteomics.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.C.L. (ljiao@ucla.edu).

Contacts

Publisher: Ben Crowe
Editor: Paul Smaglik
Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs Sales Director:

Nevin Bayouni (4978)

European Sales Manager:

Andy Douglas (4975)

UK/ RoW/ Ireland:

Nils Moeller (4953)
Irene Vigila-Atton (4944)

Scandinavia/ Spain/ Portugal:

Evelina Rubio Håkansson (4973)

Natureevents:

Sille Opstrup (4994)

France/ Switzerland:

Amelie Pequignot (4974)

Advertising Production

Manager: Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Niamh Shields

European Satellite Office

Germany/ Austria/ Italy/

The Netherlands/ Belgium:

Patrick Phelan

Tel +49 89 54 90 57 11

Fax +49 89 54 90 57 20

e-mail: p.phelan@nature.com

Reya Silao

Tel +44 (0) 20 7843 4970

e-mail: r.silao@nature.com

US Head Office, New York

345 Park Avenue South,

10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@nature.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Harakakamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami

e-mail: rasami@naturejpn.com

naturejobs

Management potential

There is rarely a foolproof method for finding the right management for biotech companies, concluded a panel at the Anglo-Nordic biotech conference last month at the London Stock Exchange. And it can be equally tricky for scientists wanting to become managers to pick a route to leadership. Geography, too, can complicate things: European companies wonder if they should consider experienced candidates from the United States, or less experienced individuals from Europe. There is a lot riding on these decisions — for individuals it's their careers, and for companies it is money from venture capitalists.

For scientists seeking management training, there are two traditional routes (see page 124). They could take some formal training, such as an MBA, a certificate programme or even just a few courses from a business school. Or they could adopt the 'just do it' attitude and plunge right into a company (assuming, of course, that they can convince someone to hire them).

At the London meeting, Anki Forsberg, a partner with the Stockholm-based venture-capital company HealthCap, explained that intangibles may be at least as important as tangibles when she considers whether HealthCap should put its money into a company. And the quality of the management is often a deciding factor.

Forsberg recommends picking companies that have chief executives she'd be comfortable inviting home for dinner. Those people aren't always the scientists whose discoveries are the basis for a start-up. "He's brilliant in the lab. Let him stay in the lab," is a refrain she uses when a scientist wants to become a chief executive of his own company, she said at the London meeting.

In the end, there are no guidelines guaranteeing success, just many possible routes and a leap of faith.

Paul Smaglik

Naturejobs editor



Contents

SPECIAL REPORT

How to become an entrepreneur p124

CAREER VIEW

Nuts & Bolts

Soft skills

Graduate Journal

Finding the balance

Movers

Marye Anne Fox p126

WWW.NATUREJOBS.COM

Career centre

Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

Being good at research isn't enough for success in the biotech business. Virginia Gewin finds out where to get the other skills you need.

Biotech is big business these days, and record numbers of PhDs are trading the ivory tower for industry. The 1,500 biotech companies in the United States alone have combined revenues exceeding \$40 billion, according to the Biotechnology Industry Organization — and they need skilled management to keep growing.

But a science education and experience in an academic lab don't provide many chances to learn about business. You may be an expert in recombinant DNA, but the business world is about making a return on investment. Although the thought of more study is less than appealing to most new PhDs, understanding business jargon and a businessperson's view of the world are critical in making the transition.

IVORY TOWER TO OFFICE TOWER

"Scientists often underestimate the importance of business skills," says Hamid Bouchikhi, chairman of the Entrepreneurship Innovation and Small Business Network at the ESSEC Business School in Paris, France. Unfortunately, good ideas don't necessarily translate into good business. "There are plenty of stories about scientists going on to lead successful companies, but nobody talks about the many companies that failed because scientists didn't make good managers," says Shreefal Mehta, research assistant professor of biotechnology management at the Rensselaer Polytechnic Institute in Troy, New York.

Indeed, leaders of this nascent industry bemoan the lack of basic skills for leadership, competitive analysis and budgeting. Not surprisingly, industry grumblings have been answered by business-school programmes located near biotech hubs such as Cambridge, Massachusetts, and San Diego, California.

"Over 95% of new PhDs stay in the life-sciences industry," says Gail Naughton, dean of San Diego State University's College of Business Administration. With funding from companies such as Invitrogen, the university started the first joint life-sciences PhD/MBA. The business models aren't as good as the scientific ideas flowing from the community, says Greg Lucier, chief executive of Invitrogen. One of the biggest problems? Scientists in industry don't understand what customers want to buy. To gain practical business experience, Invitrogen is providing programme participants with internships.

The short course is another route. The Extension Service at the University of California, San Diego offers a six-month leadership and management programme.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Employees no longer interested in a pure technology role come one day a week to learn the skills necessary to become team leaders or managers.

Even if a scientist has no wish to be an entrepreneur, understanding business will be valuable. "I can guarantee with 100% certainty that today's scientist will be involved at some level with a venture-backed company," says David Anthony, a partner in the venture-capital firm 21 Ventures in New York. Anthony has established a short course — taught at the University of Alabama in Birmingham and at the New York Academy of Sciences — and a one-year accelerated MBA programme for graduates in the life sciences, which will begin at the University of Alabama later this year. The core of the course is traditional business, replete with financing and marketing, but Anthony admits that teaching business vision is difficult. "I've found it most effective to bring in

O. STEININGER/CORBIS

outside people — entrepreneurs — to relay their own experience,” he says.

These programmes serve postdocs, but others help industry scientists switch gear to management relatively early in their careers. The Massachusetts Institute of Technology (MIT) offers the Biomedical Enterprise Program — a joint collaboration between its Sloan School of Management and the Harvard–MIT division of health sciences and technology. The programme only accepts those with substantial work experience and is attracting a wide range of people — including undergraduates seeking advice on how to set themselves up for later.

Input and support from the local business community is a hallmark of biotech-focused programmes at Simon Fraser University in Vancouver, Canada, as well as Rutgers University in Newark, New Jersey. In addition to basic business skills, including business development and good manufacturing practices, Simon Fraser offers courses, such as one on bioethics, that reflect the current needs of industry.

“Our programme is a delicate balancing act between what you need to know tomorrow and 20 years from now,” says Michael Parent, the university’s academic director of MBA programmes. The philosophy is that management cannot be taught from scratch, so the courses only take people with work experience.

Why get an MBA if you are already working in industry? “The MBA is an objective signifier of a person’s ability to work hard and get along with others,” says Parent. Mehta agrees, adding that following the technical track up the ladder from project management to group management is also a viable option. Although biotech-focused MBAs are not yet available in Europe, some of the finest business schools in the world — including ESSEC, the London Business School and IMD International in Lausanne, Switzerland — are options for those seeking core business skills that will be transferable to any industry.

BEATING THE COMPETITION

For those ready to unleash the inner entrepreneur, Britain’s Biotechnology Young Entrepreneur scheme, organized jointly by the University of Nottingham and the Biotechnology and Biological Sciences Research Council, has trained about 1,000 postgraduates in its ten-year history. Participants in the three-day competition develop business plans for imaginary biotech companies, which are judged by a panel of experts. “It’s a quick hit experiential learning for postgraduates and postdocs in the biological sciences,” says Simon Mosey, lecturer in enterprise and innovation at the University of Nottingham. For some 25 people, it was the beginning of their business. British researchers can also test the industrial waters through one-year Medici fellowships offered to scientists from the 15 participating universities.

Lack of leadership skills is holding back biotech in mainland Europe too. Carl Johan Sundberg, co-founder of the Karolinska Investment Fund and acting president of the Stockholm School of Entrepreneurship, fills the void with short courses on topics ranging from creating a business idea to starting



Networking is essential says Carl Johan Sundberg (left). Evening classes at the Karolinska provide the opportunity.



a venture. “At Stockholm, we’re mostly focused on earlier-stage people, but want to train those people with good ideas, such as ideas to patent, but don’t know how to proceed,” he says. He hopes to turn the short courses into a MBA degree in the near future.

Large drug companies remain the biggest life-sciences employers in Europe and offer an alternative route to biotech. “Big pharma is a good training ground and way into biotech industry,” says Michael Saunders, chief operating officer at 4 AZA Bioscience in Leuven, Belgium. Saunders’ time spent working in pharma inspired him to get an MBA.

GET NETWORKING

Bouchikhi describes a vicious circle of a management vacuum discouraging venture capitalists. “Venture capitalists complain that projects are not mature in terms of marketing or financial planning, which keeps them from investing,” he says. The Paris Biotech Incubator, in which ESSEC is involved, is one of a number of start-up support schemes existing in Europe’s biotech hubs. These are most suited to people with ideas at an advanced state who need to network for the few available start-up dollars. And by offering internships as entry into the business world, start-ups can attract students who believe in the company’s products and will continue with it in the hope of cashing in on its future success.

Networking is a prime motive behind US industry organizations such as the Massachusetts Biotechnology Council or San Diego’s Biocom. These have a vested interest in maintaining links with potential employees. A predicted 64% increase in demand for biotech workers in Massachusetts alone over the next decade will require researchers and clinicians as well as manufacturers. MassBioEd, the Massachusetts Biotechnology Council’s non-profit foundation, is actively educating scientists in the relevant business areas through three-day short courses.

There are many paths from lab bench towards the boardroom. All offer an individual their own way to combine scientific sense with business acumen.

Virginia Gewin is a freelance science writer based in Portland, Oregon.

Web links

San Diego State University Joint PhD MBA program
www-rohan.sdsu.edu/~cba/grad/phd.html#
 UAB Business School MBA
www.business.uab.edu/mbaforscientists
 MIT Biomedical Enterprise Program
bep.mit.edu/index.php
 Stockholm School of Entrepreneurship
www.sses.se
 Midlands Medici Fellowship Program
www.midlandsmedici.org/
 Biotechnology Yes
www.biotechnologyyes.co.uk/overview.html
 UC San Diego LAMP
extension.ucsd.edu/lamp
 Essec Business School
www.essec.edu/home
 MassBioEd survey link
www.massbio.org/massbioed/pdf/web_sum_tns_findings_final.pdf

GRADUATE JOURNAL

Finding the balance

Scientific work has two opposing demands: the first is the development of sound hypotheses, the second is testing them by hands-on work in the lab. Sometimes I find it hard to keep the balance right.

A prerequisite for a good theory that can subsequently be tested in the lab is a broad knowledge of the field. So I wade through piles of articles and books, follow up contradictory aspects and take a stance. New models must be imagined, including their implications. All this is so intriguing but so time-consuming that I'm regularly getting lost in it. And this is only half the job.

Going back to the bench to investigate the theory takes even more time and energy. Things are often not quite as straightforward as imagined. Nature always seems to work counter to human imagination. So testing a hypothesis usually results in adapting it, which requires more testing — and so on.

At some point I must draw a conclusion and I have never felt really confident in doing so. This is not because my presumptions or results are weak. But there are always more papers to read and more experiments to do. And no one has tried any of this before. Still, that's the thrill of science. It's always new. ■

Tobias Langenhan is a first-year graduate student in neuroscience at the University of Oxford, UK.

NUTS & BOLTS

Soft skills

You're a skilled researcher, steeped in the rigours of scientific enquiry and successful at your work. Your career, one might assume, will be smooth sailing? Not necessarily.

Even for a career anchored in scientific excellence, mastery of a discipline only accounts for a third of what is needed for success. You will also need organizational savvy and personal effectiveness.

Organizational savvy covers general business skills as well as the ability to navigate the politics of a particular work culture. Personal effectiveness covers the myriad 'soft' skills that, left unpractised, can derail the career of even the most accomplished professional.

If you have ever lamented your lack of budgeting experience, felt inept when negotiating a pay rise or found yourself



With Deb Koen
Careers consultant

in the political cross-fire of lab heads, then organizational savvy is the place to focus your energy. If, on the other hand, you are perceived as being aloof, or you have doubts about the strength of your communication skills, then personal effectiveness is where you should focus.

Start by getting feedback from people who work closely with you and are familiar with your style and contributions because they can serve as honest critics. If someone suggests that you work on your interpersonal skills, ask for specific examples to help

frame a 'before' and 'after' picture. A final question to ask is: "What should I start doing, stop doing and keep on doing?" When you have collected feedback from a number of people, use your scientific know-how to look for patterns in the responses.

Next, devise a strategy for adopting new behaviour. Role models are everywhere; start by observing a few. Incorporate some of their good practice into your own repertoire. Supplement these efforts with relevant coursework such as classes in active listening or budgeting for non-financial professionals. As you study and adopt new skills, be sure to close the loop by getting back to your original feedback providers. Continue to monitor your progress, and celebrate your successes. ■

Deb Koen is vice-president of Career Development Services and a columnist for The Wall Street Journal's CareerJournal.com.

MOVERS

Marye Anne Fox, chancellor, University of California, San Diego



Marye Anne Fox, who was inaugurated as chancellor of the University of California, San Diego (UCSD) in March, knows flexibility and relationship-building can have unexpected pay-offs.

Fox originally wanted to teach chemistry and physics to high-school students. But when her first husband opted to do his medical residency in a small town in her home state of Ohio, Fox found no teaching jobs. So she began graduate work.

And when her husband got drafted and assigned to Andrews Airforce Base near Washington DC, she landed a

postdoc position at the University of Maryland in College Park. When his military stint ended, her husband could set up in medical practice anywhere, and was willing to follow Fox wherever her next job took her.

This turned out to be the University of Texas, Austin, where she became the first female assistant professor. She found she relished administrative tasks as much as research. So a chance meeting with former professor Norman Hackerman, who maintained an office in Fox's building, opened new doors. "He really served as a mentor to me, and suggested my name for a number of committees," Fox says.

Fox says she never experienced any tangible discrimination in the early stages of her career. But she was often the only woman on local and national committees, which sensitized her to issues of discrimination and the under-representation of women and minorities

in science. Fox thinks one answer to under-representation is to make people who bring in new faculty members and staff accountable. As part of staff members' yearly appraisal at UCSD, she will ask what efforts they've made to recruit women and minorities. She stresses that this is not an affirmative action approach, but an attempt to ensure that women and minorities are at least in the pool of candidates.

Although discouraged that the number of women receiving tenure-track posts is far fewer than the number receiving PhDs, she still tells women that "there's never been a more exciting time in science" and encourages them to try, despite the obstacles.

These may include raising a family while advancing one's career. But Fox says she is proof it can be done, while maintaining a balanced life. "I'm no superwoman, by any means," she says. ■

CV

1998–2004: Chancellor and distinguished university professor of chemistry, North Carolina State University, Raleigh.

1976–98: University of Texas, Austin (advancing from assistant professor of organic chemistry to vice-president for research).

Meat

A growth industry.

Paul McAuley

We certainly don't call any of our clients 'The Meat', or 'Pork Chop #1'. That's just tabloid nonsense. And while we're skewering misconceptions, the job isn't as glamorous as you might think. Although I go to club and restaurant openings, film premières, first nights, fashion shows and stay in first-class hotels all around the world, I'm not there to enjoy myself. I'm there to prevent any live cells from my clients' bodies falling into the hands of meatleggers.

Sweep, security and clean-up: that's what the job is about. First, my team sweeps the place before the client arrives, everything from running background checks on staff to inspecting the restrooms. Restrooms are where the clients are most vulnerable, of course. We run fibre optic cameras down the pipework, looking for traps and filters; we check for microscopic rasps designed to trap a few skin cells by fogging the place with an aerosol of tailored bacteria and using ultra-violet light to spot any unusual clumping. We inspect seating, too; tableware at restaurants; glasses at bars; cocktail napkins — we have detailed checklists for every type of venue.

After the client arrives, we run an extra layer of security. No known meatlegger gets within a hundred yards of any of my perimeters, but they'll bribe staff or plant some innocent-looking old lady in the crowd, and it's not unknown for some minor celebrity in need of quick cash to try to snag a few cells from an A-lister. I have a database of known stooges and my people keep a lookout for abnormal behaviour, I can't tell you any more than that. Afterwards, there's the clean-up, which is the most routine but most important part of the job, and which I always supervise personally. Some cleaners ride shotgun on their clients on the way back from an event, but I like to make sure that the venue is sterilized more thoroughly than any operating theatre. The bodyguards can look after the client in transit, and besides, once

they're in their limo, my clients are protected by a Class Four biohazard containment environment. Not even a virus can get in or out.

Hotels? That's a whole book right there. Any place a client of mine would use has its own cleaning protocol, but I like to think I add my special magic to the mix.

Like a lot of cleaners, I started out in public

As soon as anyone managed to get a viable scrap of tissue, that was it. The meat was out there. The only way to stop it was to bust the places where it was grown.

Dangerous? Not really. The meat trade is too specialized to interest professional criminals, although quite a few are customers; one crime boss likes to serve the meat of his ene-

mies with his special sauce. Politicians and business-people also enjoy revenge feasts, but the fans are the backbone of the trade. These days, you aren't a hardcore tru-fan unless you've partaken of the flesh of your hero. It's the ultimate form of possession, and I don't suppose I need to point out the parallels with Christian communion. No, that's just an urban myth lifted from some cheesy bestseller. In order to clone tissue, you need to start with live cells, or at least a live nucleus, and after 2,000 years ... exactly.

Cloned babies? Another myth. It's very difficult to turn a somatic cell into an embryo, and even harder to bring it to term. Far easier to grow sheets of epidermis or muscle. I guess the oddest case I dealt with was the meatlegger who cloned himself. All he ate was his own meat. I guess you could say he was really into self-sufficiency.

I don't think the meat trade is going to die out any time soon. Most clone lines have been wiped out, and

people like me do their best to make sure that the meatleggers have a hard time getting fresh ones started, but now there's this new thing. These nanotech makers. Pretty soon the meatleggers won't need live cells, just a DNA sequence, and I've heard that these makers can build an entire body from scratch.

As long as people keep finding twisted uses for new technology, there'll always be a need for people like me, cleaning up the mess.

Paul McAuley has worked as a research biologist at various universities including Oxford and the University of California, Los Angeles, and for six years was a lecturer in plant science at St Andrews University. His latest novel is White Devils (Simon & Schuster).



health, running DNA analyses in a forensic laboratory. That was ten years ago, when the meat trade was at its height. We were processing 10,000 samples a day. Most were fakes. 'Princess Di' for instance, was originally a basal-cell carcinoma excised from a 58-year-old Albanian woman, but it didn't stop the meatleggers moving 20 tonnes of product. Then fans started doing their own DNA analyses, and growing their own supplies. Once someone has started a cloned cell line, anyone with an incubator, access to a few common biochemicals, and basic knowledge about cell culture can keep it going indefinitely. By the time I joined one of the vat-busting teams, most of the meat we were chasing was 100% genuine cloned celebrity.